Università di Pisa
Dipartimento di Fisica

School of graduate studies
Galileo Galilei
XIX ciclo

Ph.D. Thesis

Search for the Standard Model Higgs decaying into tau–jet final state with the ATLAS experiment at the LHC

Candidate

Francesca Sarri

Supervisor

Chiar.mo Prof. **Vincenzo Cavasinni**

Contents

Introduction

The Large Hadron Collider (LHC, at CERN) will be the next milestone in our quest to understand physics at the most fundamental level. The design energy is one order of magnitude larger than that of the most powerful existing hadron collider (the TeVatron at Fermilab). Both experimental and theoretical physicists hope and expect that the LHC will reveal a lot of new physics.

The most important goal of the two general-purpose experiments at LHC (ATLAS and CMS) is the unraveling of the Electroweak symmetry breaking mechanism. In fact, both the Standard Model (SM) of the Elementary Particles and its supersymmetric extensions need a mechanism that breaks the gauge symmetry of the lagrangian, thus giving mass to fermions and bosons. Both in the Standard Model and in its minimal extensions, one or several new bosons are expected as manifestations of the Higgs field, responsible for this symmetry breaking.

Both ATLAS and CMS have dedicated many studies to the discovery of the Higgs boson(s). The design of both experiments has in large part been driven by requirements to achieve that goal.

The theoretical predictions foresee that the dominant Higgs boson decay mode will be $H \to b\bar{b}$ followed by $H \to \tau^+\tau^-$, in most of the allowed mass range. In the framework of the Standard Model, these decay modes dominate if the Higgs mass is less than 130 GeV. Also in the Minimal Supersymmetric extension of the Standard Model (MSSM), these decay channels play a significant role, both for the light Higgs boson h and for the heavy neutral bosons A and H. The detection of the $H \to b\bar{b}$ decay channel is not feasible because of the huge QCD background that will be present at LHC. Because of the large cross section for direct production of b–jets, it is impossible to extract the Higgs signal from the background.

The second most favorite decay channel in this mass region is $H \to \tau^+\tau^-$ with a branching ratio (BR) about a factor 8 lower. Despite of the lower cross section this decay channel seems the most promising for the Higgs detection especially when the associated production mechanism is considered.

The vector boson fusion (VBF) production mechanism, in which the Higgs boson is accompanied by two energetic forward jets, thanks to its peculiar final state topology allows a strong rejection against background, and is an important channel for Higgs discovery and for the measurement of the Higgs boson coupling to fermions, since beyond the $b\bar{b}$ decay mode, no other direct fermion decay mode is accessible at LHC.

Studies have been done in the past on VBF $H \rightarrow \tau^+\tau^-$. However, only those channels in which either both τ's or at least one τ decays leptonically have been analyzed. The reason is the huge QCD background dominating by several order of magnitudes any low cross section process involving fully hadronic final states. The difficulties arise already at the trigger level. Due to the fact that the Level 2 Trigger (LVL2) is not very effective in the rejection of full hadronic Level 1 (LVL1) accepted events, the LVL1 bandwidth for multijet trigger has to be kept low by keeping very high jet $\mathrm{E_T}$ trigger thresholds. On the other hand, the τ-jets have a different topology compared with ordinary QCD jets. They are made of one or three charged hadrons plus some neutrals and they are narrower than the others. These characteristics make possible a τ–jet tag already at LVL1. Since the single tau trigger is expected to be exposed to the large QCD jets background, the ATLAS trigger menu foresees also a tau trigger in combination with other signatures, like $\mathrm{E_T^{miss}}$ or another τ in the event. Moreover the algorithms of the ATLAS offline τ reconstruction provide a very good rejection against ordinary QCD jets.

In this thesis we focus on the feasibility of the measurement of VBF $H \rightarrow \tau_h^+\tau_h^- + \mathrm{E_T^{miss}}$ in fully hadronic final states.

The first chapter provides an overview of the Higgs mechanism, both from the theoretical and the experimental point of view. A brief review of the theoretical aspects of the standard model is followed by a summary of the most important theoretical constraints on the SM Higgs mass. The current experimental limits are then discussed. In the last part of the chapter, we consider the SM Higgs discovery potential with the ATLAS detector.

The second chapter presents details of the ATLAS detector. After a brief description of the LHC, we describe the inner detector, the calorimeters, the muon spectrometer and the trigger system.

The third and fourth chapters are dedicated respectively to the trigger and to the reconstruction in the ATLAS detector of the hadronic τ and of the $\mathrm{E_T^{miss}}$.

Finally, in the fifth chapter a possible discovering of the Higgs boson through the measurement of $qqH \rightarrow qq\tau_h^+\tau_h^- + \mathrm{E_T^{miss}}$ channel is discussed.

Chapter 1

The Higgs Boson: theoretical predictions and experimental challenges

The Electroweak Symmetry Breaking via the introduction of a scalar complex doublet in the Standard Model (SM) Lagrangian provides a way both to give masses to fermions and bosons and to cancel divergences in the WW scattering amplitude. The doublet manifests itself with the presence of a new particle, the Higgs boson, whose mass is not predicted by the model. The aim of the present work is to verify the possibility of measurements in the Higgs sector with the use of full hadronic final states. The present chapter provides a theoretical overview of the SM Higgs sector, together with the present experimental limits and the prospects for the Higgs discovery with the ATLAS apparatus.

Introduction

The requirement of gauge invariance under $SU(2) \otimes U(1)$ transformations of the SM ([1],[2],[3]) massless Lagrangian forbids the introduction of any direct mass term for the fermions and gauge bosons, which is in contrast with the experimental evidence. Therefore, a mechanism to provide masses to the particles is needed. By introducing a complex scalar doublet together with a quartic self interaction potential, the vacuum expectation value (VEV) for this field can be made different from zero due to spontaneous symmetry breaking ([4],[5],[6]). The presence of this VEV gives masses to the SM fermions and to three gauge bosons each of which absorbs a Goldstone boson corresponding to one of the three broken symmetry generators thus acquiring

	doublets L	singlets R
Leptons	$\binom{e}{\nu_e}_L$, $\binom{\mu}{\nu_\mu}_L$, $\binom{\tau}{\nu_\tau}_L$	e_R , μ_R , τ_R
	$\lvert\overrightarrow{T}\rvert = \frac{1}{2}$ $\quad Y = \frac{1}{2}$	$\lvert\overrightarrow{T}\rvert = 0$ $\quad Y = -1$
Quarks	$\binom{u}{d}_L$, $\binom{c}{s}_L$, $\binom{t}{b}_L$	u_R , d_R , c_R , s_R , t_R , b_R
	$\lvert\overrightarrow{T}\rvert = \frac{1}{2}$ $\quad Y = \frac{1}{6}$	$\lvert\overrightarrow{T}\rvert = 0$ $\quad Y(u_R, c_R, t_r) = \frac{2}{3}$
		$Y(d_R, s_R, b_r) = -\frac{1}{3}$

Table 1.1: *Doublets and singlets of the weak isospin for leptons and quarks with the isospin and hypercharge quantum numbers.*

a longitudinal mode. The remaining scalar field gives rise to a new particle, the Higgs boson, whose mass is an unknown (input) parameter of the model. Theoretical upper bounds on the Higgs mass can be derived by the constraint of unitarity of the scattering amplitudes of longitudinally polarized W bosons, while lower bounds can be obtained by requiring the stability of the vacuum. The LEP experiments have experimentally excluded the existence of a SM Higgs boson with masses lower than 114.4 GeV [1]. The ATLAS experiment will search for the Higgs boson from the LEP experimental limits up to energies of order $\simeq 1$ TeV.

The present chapter provides a brief theoretical overview of the Electroweak Symmetry Breaking mechanism in the SM. After a description of the Higgs sector of the SM and a short review of the experimental limits on the Higgs mass, the production and decay channels useful for the Higgs discovery with ATLAS are discussed.

The aim is mainly to clarify the context in which the analysis of the last chapter of the present work have been done. The measurement of $\tau^+\tau^-$ final states from the Higgs decays is extremely important for the $\tau^+\tau^- H$ Yukawa coupling measurement as well as for the Higgs discovery.

[1] We use the natural units ($\hbar = c = 1$)

1.1 The Standard Model Higgs Boson

The requirement of gauge invariance under any transformation of the group $SU(2) \otimes U(1)$ modifies the free Lagrangian for the fermions:

$$\mathcal{L} = i\bar{\psi}_R \gamma_\mu \partial^\mu \ \psi_R + i\bar{\psi}_L \gamma_\mu \partial^\mu \ \psi_L \ , \tag{1.1}$$

into the following:

$$\mathcal{L} = -\frac{1}{4}\mathbf{W}_{\mu\nu}\mathbf{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \bar{\psi}_L i\not{D} \ \psi_L + \bar{\psi}_R i\not{D} \ \psi_R \ , \tag{1.2}$$

where

$$\mathbf{W}_{\mu\nu} = \partial_\mu \mathbf{W}_\nu - \partial_\nu \mathbf{W}_\mu + g(\mathbf{W}_\mu \times \mathbf{W}_\nu) \ , \tag{1.3}$$

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \ , \tag{1.4}$$

$$\not{D} = \gamma^\mu D_\mu = \gamma^\mu (\partial_\mu - igW_\mu^i \frac{\tau_i}{2} - ig' B_\mu \frac{Y}{2}) \psi_L \ . \tag{1.5}$$

$\mathbf{W}_\mu$ and B_μ are gauge boson fields introduced to preserve respectively the $SU(2)$ and the $U(1)$ invariance, τ_i are the generators of $SU(2)$, g and g' are the coupling constants of the gauge field to the fermions. Any mass term for the fermions would involve expressions like

$$\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L \ , \tag{1.6}$$

which are not gauge invariant. A mass term for the gauge bosons also violates the gauge invariance. An elegant way to give masses both to the fermions and the bosons is provided by the Higgs mechanism. If one introduces a new complex scalar $SU(2)$ doublet

$$\varphi = \begin{pmatrix} \varphi_+ \\ \varphi_0 \end{pmatrix} \tag{1.7}$$

in the Lagrangian, the most general terms allowed by the gauge invariance and by the renormalizability are:

$$\mathcal{L}_H = (D_\mu \varphi)^\dagger (D^\mu \varphi) - V(\varphi^\dagger \varphi) + \chi \ (\varphi \bar{\psi}_L \ \psi_R + h.c.) \ , \tag{1.8}$$

where the potential V is given by the following expression:

$$V(\varphi^\dagger \varphi) = \mu^2 \varphi^\dagger \varphi + \lambda (\varphi^\dagger \varphi)^2 \ . \tag{1.9}$$

The minimum of the potential in the Eq. (1.9) (vacuum expectation value, or VEV of the Higgs field) is in $\varphi = 0$ if $\mu^2 > 0$ (conventionally $\lambda > 0$), but if μ^2 is negative, then the minima of the potential lie on a circle of radius

$$|\varphi|^2 = -\frac{\mu^2}{2\lambda} \ . \tag{1.10}$$

The Higgs field can be expanded around one of the minima ($\varphi = \bar{\varphi} + \varphi'$). A gauge choice can be made so that any component of the field different from $\bar{\varphi}$ is zero (unitary gauge). The SM Lagrangian becomes

$$\begin{aligned} \mathcal{L} \ = \ & -\frac{1}{4}\mathbf{W}_{\mu\nu}\mathbf{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \bar{\psi}_L i\not{D}\ \psi_L + \bar{\psi}_R i\not{D}\ \psi_R \\ & + (D_\mu\varphi')^\dagger(D_\mu\varphi') + \bar{\varphi}^\dagger(igW_\mu^i\frac{\tau_i}{2} + i\frac{g'}{2}B_\mu)(-igW_j^\mu\frac{\tau_j}{2} - i\frac{g'}{2}B^\mu)\bar{\varphi} \\ & - V(|\bar{\varphi} + \varphi'|^2) + \chi(\varphi'\bar{\psi}_L\psi_R + h.c.) + \chi(\bar{\varphi}\bar{\psi}_L\psi_R + h.c.) \ . \end{aligned} \tag{1.11}$$

The last term is a mass term for the fermions. The Higgs coupling to the fermions χ is therefore proportional to the fermion mass, while the matrix mass for the vector bosons is

$$\frac{M^2}{2} = \frac{\bar{\varphi}^2}{4}\begin{pmatrix} g^2 & 0 & 0 & 0 \\ 0 & g^2 & 0 & 0 \\ 0 & 0 & g^2 & -gg' \\ 0 & 0 & -gg' & g'^2 \end{pmatrix} \ . \tag{1.12}$$

The diagonalization of the matrix involves a rotation of an angle θ_W (the Weinberg angle) in the $W_0 - B$ space. The following relations between the Weinberg angle and the coupling constants g and g' exist:

$$\sin\theta_W = \frac{g'}{\sqrt{g^2 + g'^2}} \ , \ \cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}} \ . \tag{1.13}$$

The mass eigenstates are:

1. $W^\pm$ with bare mass $M_W^2 = \frac{g^2\bar{\varphi}^2}{2}$;
2. Z_0 with bare mass $M_Z^2 = \frac{g^2+g'^2}{2}\bar{\varphi}^2$;
3. A with bare mass 0.

The value of the VEV $\bar{\varphi}$ can be extracted by its relation with the W mass, and it turns out to be $\bar{\varphi} = 246$ GeV.

The degrees of freedom for the longitudinal polarizations of the vector bosons come from those removed by the gauge choice. The value of $\sin^2\theta_W$

is extracted from Z–pole observations and neutral current processes ([7],[8]), and depends on renormalization prescriptions. There are several different schemes that lead to slightly different values. In the modified minimal subtraction scheme ([9],[10]), the most recent value obtained by combining all the electroweak measurements of several experiment is [11]

$$\sin^2\theta_W \simeq 0.23120 \pm 0.00015 \ . \tag{1.14}$$

1.2 The Higgs Couplings

At the tree level, the coupling between the Higgs boson and the fermions (χ) is determined by the last but one term of eq. 1.11. Since the masses of the fermions are equal to $\chi\bar{\varphi}$, and the value of the Higgs field minimum $\bar{\varphi}$ can be expressed in terms of the W boson mass, the coupling to the fermions turns out to be

$$\chi = g_{ffH} = \frac{g m_f}{2M_W} \ . \tag{1.15}$$

Therefore, the coupling is proportional to the fermion mass. The partial width of the Higgs boson into any fermion channel is ([4],[5],[6])

$$\Gamma(H \to \bar{f}f) = \frac{N_c g^2 m_f^2}{32\pi M_W^2}\beta^3 m_H \ , \tag{1.16}$$

where N_c is the color factor (1 for leptons, 3 for quarks) and $\beta^2 = 1 - 4m_f^2/m_H^2$.

The (tree level) partial widths for the Higgs boson decay into on–shell vector bosons are:

$$\Gamma(H \to W^+W^-) = \frac{g^2 m_H^3}{64\pi M_W^2}\sqrt{1-x_W}\left(1 - x_W + \frac{3}{4}x_W^2\right) \ , \tag{1.17}$$

$$\Gamma(H \to ZZ) = \frac{g^2 m_H^3}{128\pi M_W^2}\sqrt{1-x_Z}\left(1 - x_Z + \frac{3}{4}x_Z^2\right) \ , \tag{1.18}$$

with $x_i (i = W, Z) = 4M_i^2/m_H^2$. If $m_H \gg M_i$, the terms involving x_i tend to one and can be therefore dropped in the above expression. As a consequence, the Higgs partial width for decays into vector bosons increases as m_H^3 with the Higgs mass. An easy–to–remember rule is

$$\Gamma(H \to VV) \simeq \frac{1}{2}m_H^3 \ , \tag{1.19}$$

where both the partial width and the Higgs mass are expressed in TeV. In the intermediate Higgs mass region (130 GeV $< m_H <$ 200 GeV), the off–shell decays into vector bosons play also an important role.

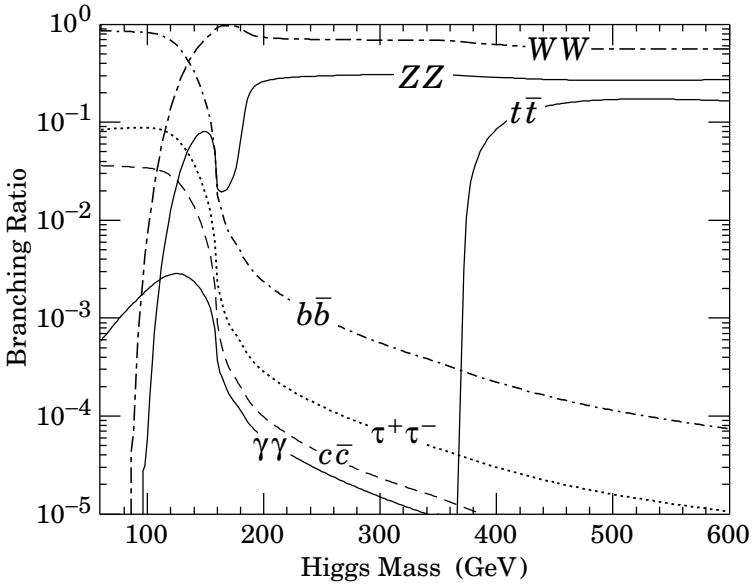

Figure 1.1: *Branching ratios of the SM Higgs boson as a function of its mass.*

Higgs decays into gluon–gluon and $\gamma\gamma$ are mediated by fermionic loops. Since the Higgs coupling to the fermions is proportional to the mass, the main contribution comes from the top quark and (in the latter) from W loops.

Figure 1.1 shows the ratios of the Higgs into different final states as a function of the Higgs mass. In the low mass region ($m_H < 130$ GeV) the most important decay channels are $H \to b\bar{b}$ and $H \to \tau^+\tau^-$. A much lower BR is present for the $\gamma\gamma$ decay, but the clear signature and the relatively low background make this decay useful for the Higgs discovery at LHC, where the discovery in the $b\bar{b}$ channel is difficult due to the huge QCD di–jet background (the cross section for inclusive $b\bar{b}$ production is $\sigma_{b\bar{b}} \simeq 100$ μb). In the rest of the mass range the decay into vector bosons provides a clear signature to extract the Higgs signal from the background.

1.3 Limits on the Higgs Mass

The Higgs mass cannot be theoretically predicted. Nevertheless, theoretical bounds on m_H can be set from internal consistency conditions and extrapolations of the model to high energies. Moreover, the LEP experiments in particular have set a lower limit on the Higgs mass by direct searches, and upper and lower bounds, coming from consistency measurements of many

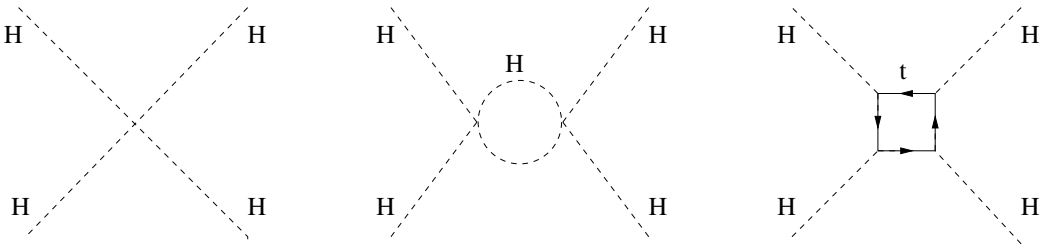

Figure 1.2: *Feynman diagrams that mainly contribute to the Higgs quartic coupling.*

parameters of the SM that have been measured.

1.3.1 Theoretical Limits

In the absence of a Higgs boson, the amplitude for the elastic scattering of massive W bosons grows indefinitely with the energy for longitudinal polarized particles, as a consequence of the linear rise of the longitudinal W_L wave function $\epsilon_L = (p, 0, 0, E)/M_W$. On the other hand, unitarity requires elastic–scattering amplitudes of partial waves J to be bounded by $\mathfrak{Re}A_J \leq 1/2$. Applied to the asymptotic S–wave amplitude $A_0 = G_F s/8\pi\sqrt{2}$ of the isospin–zero channel $2W_LW_L + Z_LZ_L$, the bound

$$s \leq 4\pi\sqrt{2}/G_F \simeq (1.2 \text{ TeV})^2 \ , \tag{1.20}$$

on the center of mass (c.m.) energy $\sqrt{s}$ can be derived [12] for the validity of a theory of weakly coupled massive gauge bosons.

The increase of the scattering amplitude can be damped by Higgs boson exchange. From the asymptotic expansion of the elastic W_LW_L S–wave scattering amplitude including W and Higgs exchanges,

$$A(W_LW_L \to W_LW_L) \to -G_F M_H^2/4\sqrt{2}\pi \ , \tag{1.21}$$

it follows ([13] and references therein) that

$$M_H^2 \leq 2\sqrt{2}\pi/G_F \simeq (850 \text{ GeV})^2 \ . \tag{1.22}$$

Within the canonical formulation of the SM, internal consistency therefore requires $M_H < 1$ TeV.

A further upper bound can be derived considering the evolution with the field strength of the quartic coupling of the Higgs λ. The main contributions to λ are depicted in Figure1.2. The Higgs loop itself gives rise to an indefinite increase of the coupling, while the top contribution drives the coupling to

Λ	m_H
1 TeV	55 GeV $\leq m_H \leq$ 700 GeV
10^{19} GeV	130 GeV $\leq m_H \leq$ 190 GeV

Table 1.2: *Higgs mass bounds for two values of the cut–off* Λ

smaller values with increasing top mass, eventually below zero. The upper bound comes from the request that λ becomes strong at some energy scale Λ up to which the SM is assumed to be valid. The evolution of λ with the energy μ^2 is described (for moderate top masses) by

$$\lambda(\mu^2) = \frac{\lambda(v^2)}{1 - \frac{3\lambda(v^2)}{8\pi} \log \frac{\mu^2}{v^2}} \ , \tag{1.23}$$

where v is the VEV and $\lambda(v^2) = m_H^2/v^2$. After the substitution of this relation in eq. 1.23, the condition $\lambda(\Lambda) < \infty$ translates in an upper bound on m_H:

$$m_H^2 \leq \frac{8\pi^2 v^2}{3 \log \frac{\Lambda^2}{v^2}} \ . \tag{1.24}$$

A lower bound on the Higgs mass can be derived from the stability of the vacuum. Since for increasing top mass λ decreases down to negative values, in such cases the ground state would be no longer stable. Thus, for a given top mass, m_H must exceed some minimal value, which depends on λ.

Thus, for any given energy scale Λ up to which the SM has to be valid, upper and lower bounds on the Higgs mass can be derived if the top mass is known. Table 1.2 shows the values on the Higgs mass for two different values of Λ. As can be seen, the mass bounds on the Higgs boson become tight if we assume the SM to be valid up to the Grand Unification Theory (GUT) scale ($10^{19} GeV$).

1.3.2 Experimental Limits

The most tight limits on the SM Higgs mass come from the LEP experiments. The four LEP Collaborations (ALEPH, DELPHI, L3 and OPAL) have collected a total of 2461 pb^{-1} of integrated luminosity in e^+e^- collisions at center of mass energies between 189 and 209 GeV. The main production process in e^+e^- at the LEP energies is the *Higgsstrahlung* process $e^+e^- \to HZ$ [14], with small contributions coming from the W or Z fusion ($e^+e^- \to WW\nu_e\bar{\nu}_e \to H\nu_e\bar{\nu}_e$ and the corresponding process with the Z fusion). The main decay channel is $H \to b\bar{b}$ (the BR is 74% if $m_H = 115$

GeV). There are also contributions from the decays into $\tau^+\tau^-$, WW^* and gg (7% each) and $c\bar{c}$ (4%). Therefore, the search for a Higgs boson is performed in 4 different signatures: the 4 jet channel ($HZ \to b\bar{b}q\bar{q}$), the missing energy channel ($b\bar{b}\nu\bar{\nu}$), the leptonic channel($b\bar{b}l^+l^-$) and the tau lepton channel ($\tau^+\tau^- q\bar{q}$). The results of the four collaborations in the four channels have been combined in a likelihood ratio analysis [15] to provide the final result.

The conclusion of the LEP analysis provides an experimental lower bound for the Higgs mass obtained by direct search:

$$m_H > 114.4 \text{ GeV} , \tag{1.25}$$

at the 95% confidence level. The ALEPH Collaboration data reported an excess of events consistent with a Higgs boson with $m_H = 115$ GeV, but the data of the other three collaborations are consistent with the background hypothesis.

Indirect indications for the Higgs data come from the electroweak fit of the SM. The LEP Electroweak Working Group has combined several measurements of SM parameters coming from LEP itself, as well as from other experiments (the TeVatron experiments CDF and D0, for example, for the top mass)[16]. Since the Higgs boson mass enters in the calculation of higher order corrections of those quantities, the Higgs boson mass can be inferred by a global fit of all the measured SM parameters. Using recent results from TeVatron top mass, Figure 1.3 [17] shows the results for the χ^2 of the fit as a function of the Higgs mass together with the band corresponding to the theoretical uncertainties. The preferred value for the Higgs mass is

$$m_H = 84^{+34}_{-26} \text{ GeV} , \tag{1.26}$$

with an upper limit at 154 GeV at the 95% of confidence level. This limit increases to 185 GeV when including the LEP-2 direct search limit of 114 GeV.

Prospects for the Higgs Discovery at the TeVatron

At present, the TeVatron (at Fermilab, Chicago, USA) is the most powerful existing hadron collider. The $p\bar{p}$ collisions occur with a center of mass energy of about 2 TeV. We briefly summarize the Higgs discovery chances of the two experiments that study this collisions, CDF [18] and D0 [19].

The main production channel for the Higgs at the TeVatron is the gluon fusion, as it will be at the LHC. Higgsstrahlung processes from a W or Z bosons have cross sections one order of magnitude lower, but they provide a distinctive signature for the final state. The Vector Boson Fusion production channel is suppressed with respect to the LHC.

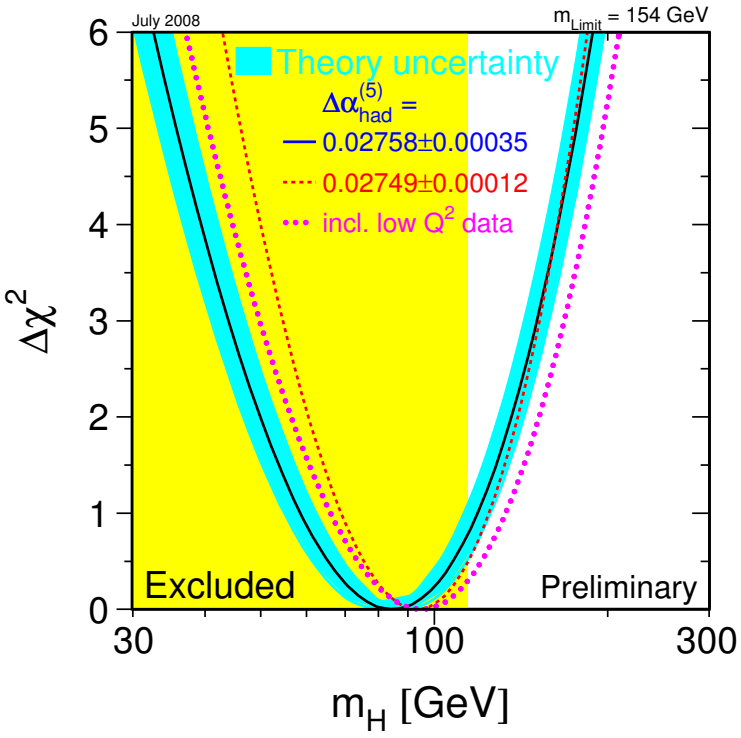

Figure 1.3: *The χ^2 of the Electroweak fit of the SM as a function of the Higgs mass. The preferred value turns out to be 84 GeV.*

The favorite decay channel for the low–mass Higgs boson is $b\bar{b}$. The direct detection in this channel is limited by the huge QCD background present in a hadron collider at the TeVatron energies. Therefore, the most important final state for the Higgs detection is $WH \to l\nu b\bar{b}$. If the Higgs mass exceeds 130 GeV, the most important decay channel for the Higgs is $H \to WW$, which provides many different signatures in the final state [20].

The key point for the Higgs discovery at TeVatron is the statistics that the experiments will be able to collect. Figure 1.4 [21] shows the integrated luminosity that is necessary to impose a 95% CL limit, for a 3σ evidence and for a 5σ discovery combining the two experiments. As can be seen, the 5σ discovery in the mass range from the LEP limits up to 130 GeV will require 10 to 30 fb^{-1} of integrated luminosity.

The plot of the luminosity so far achieved [22] can be found in Figure 1.5. The integrated luminosity expected at the end of the TeVatron operations (2009-2010) is about 8 fb^{-1} [23].

Therefore, it seems unlikely to discover a SM Higgs boson at the TeVatron. On the contrary, it will be possible to exclude or to have a 3σ evidence

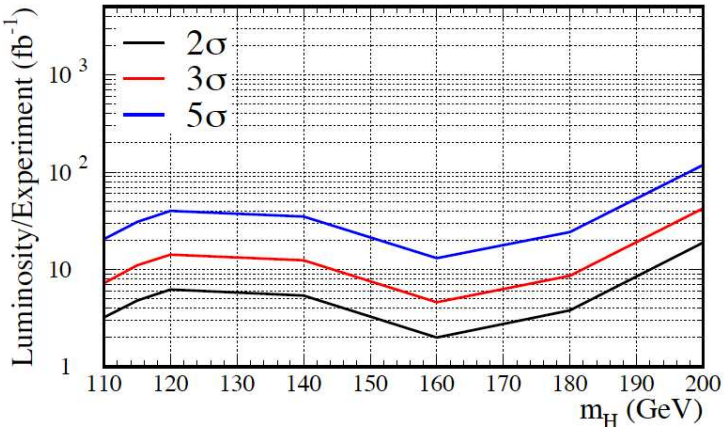

Figure 1.4: *The statistics needed by the combination of the TeVatron experiments to set a 95% CL limit, for a* 3σ *evidence and for a* 5σ *discovery as a function of the Higgs mass.*

in a wide SM Higgs mass region allowed by the precision electroweak measurements. Preliminary recent results, obtained with 3 fb^{-1} of data and combining the two TeVatron experiments, already exclude the production of a SM Higgs of 170 GeV at 95% CL [24].

1.4 The Higgs at ATLAS

One of the main task of the experiments at LHC is the SM Higgs search. It will be possible to explore the whole mass range, from the LEP limit up to the TeV scale [25]. For a description of the LHC and of the ATLAS detector, we refer to the next chapter of this work (and references therein).

The main production channels for the Higgs boson at a pp collider with $\sqrt{s} = 14$ TeV are shown in Figure 1.6. Although the largest cross section is that of the $gg \to H$ production on the whole mass range, it is often convenient to consider associated production channels. The associated production of two light quarks via Vector Boson Fusion (VBF production) is the next to leading process for the Higgs boson production. Nevertheless, the peculiar final state [26] with two hard jets in the forward and backward region of the detector (coming from the fragmentation of the two quarks after the vector bosons radiation) provides a distinctive signature that can be efficiently used to disentangle the signal from the background. Similar topological characteristics can be exploited in the other associated production channels.

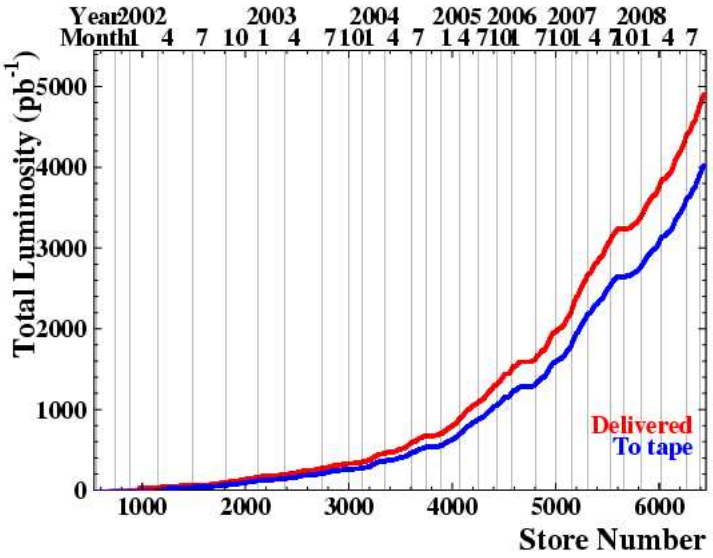

Figure 1.5: *Luminosity achieved by the TeVatron up to 2008.*

1.4.1 The Low Mass Region

If $m_H < 130$ GeV, the most important decay channel for the Higgs is $H \to b\bar{b}$ (Figure 1.1). Because of the huge QCD background ($\sigma_{b\bar{b}} \simeq$ 100 μb), the detection of the Higgs boson is not feasible in the inclusive production. Therefore associated productions can be used to gain additional rejection. This can be obtained considering final states such as $t\bar{t}H$. If a semileptonic decay is requested for one of the two top quarks, the trigger requirement is satisfied by a high p_T lepton in the final state. The complexity of the final state ($t\bar{t}H \to l\nu 2j4b$) is useful to reduce the main backgrounds to this channel ($t\bar{t}b\bar{b}$ and $t\bar{t}Z$ production).

The $H \to \gamma\gamma$ decay channel is also interesting in this mass range. The performance of the electromagnetic calorimeter is the key point to see the signal, which BR is below 1%. Therefore, most of the design criteria for the LAr calorimeter (see next chapter) have been driven by the requirements of this channel. The irreducible background comes from the direct production of $\gamma\gamma$, while the main reducible background comes from jj and $j\gamma$ final states where the jets have been misidentified as photons. This can occour especially if the jet is composed by one leading π_0 and a number of soft hadrons. The rejection against these jets requires high angular granularity to distinguish the two photons coming from the π_0 decay. This is obtained by the combination of the high granularity presampler and first strip layer of the ATLAS electromagnetic (EM) calorimeters.

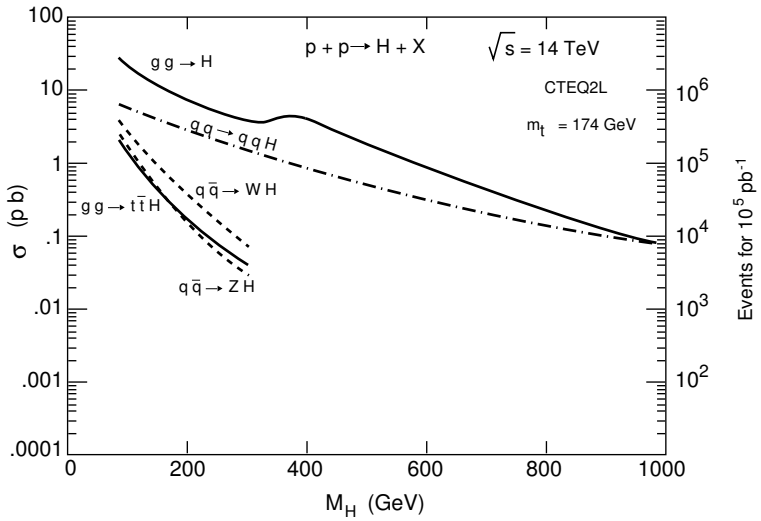

Figure 1.6: *Higgs production cross sections at LHC for different production channels as a function of the Higgs mass.*

The invariant mass distribution of the two photons is shown in Figure 1.7 on the left.

The most important contribution for the Higgs discovery in the 110 GeV $< m_H <$ 130 GeV mass range comes from the $H \to \tau\tau$ decay where the Higgs has been produced by VBF ($qq \to WW(ZZ)qq \to qqH$) [27]. The presence of the two hard forward jets together with the reduced hadronic activity between them due to the lack of color exchange between quarks in the initial state, are exploited to reduce the background. In the analysis in [27] both the $\tau\tau \to ll + \mathrm{E_T^{miss}}$ and the $\tau\tau \to l\tau_h + \mathrm{E_T^{miss}}$ final states have been considered. The main sources of background are the $Z + nj$ production and, less important, the $t\bar{t} + nj$ production. The results of this channel are reported in Table 1.3.

Recently also the $\tau\tau \to \tau_h\tau_h + \mathrm{E_T^{miss}}$ final state has been investigated [28].

M_H (GeV)	110	120	130	140	150
Signal	36.2	32.3	24.4	16.5	7.4
Background	65.4	21.4	12.2	9.0	7.7
Statistical Significance	5.6	6.5	6.5	5.8	4.3

Table 1.3: *Signal and background events expected after 30 fb^{-1} in the VBF $H \to \tau\tau$ channel.*

In this thesis new results for the fully leptonic and semileptonic final states

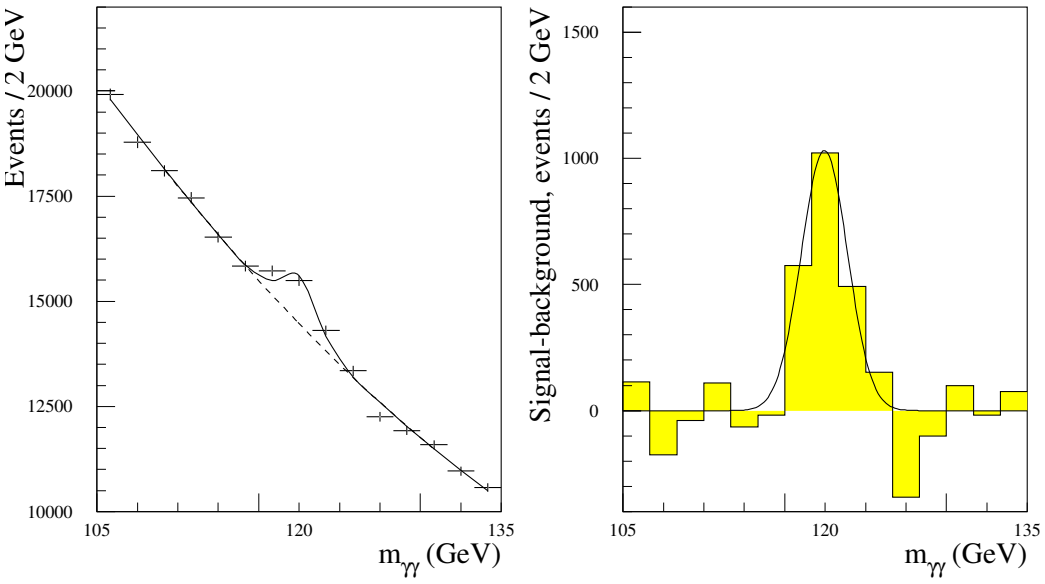

Figure 1.7: *Invariant mass distribution after the kinematical selection of the $H \to \gamma\gamma$ analysis before (left) and after (right) the background subtraction. The Higgs mass is 120 GeV and the integrated luminosity is 100 fb^{-1}.*

will be shown, together with the study of the fully hadronic final state.

1.4.2 The Intermediate Mass Region

If 130 GeV $< m_H <$ 200 GeV, the Higgs decays mostly into vector bosons. The most important channel is $H \to WW$, with the Higgs produced via Vector Boson Fusion [27]. The final states analyzed are both the semileptonic final state $WW \to ljj + \mathrm{E}_{\mathrm{T}}^{miss}$ and the full leptonic final state $WW \to ll + \mathrm{E}_{\mathrm{T}}^{miss}$, the latter being by far the most important one. The use of the VBF production features with the two hard jets in the forward and backward regions usually identified with a specific forward tagging algorithm, and the lack of hadronic activity in the central region indicated by applying a veto for jet with transverse energy above some threshold, allows rejection of the background, most importantly from $t\bar{t}$ production. Figure 1.8 shows the transverse mass distribution for the $H \to e\mu + \mathrm{E}_{\mathrm{T}}^{miss}$ final state. As can be seen, the signal is well above the background, which is mainly composed of $t\bar{t}$.

The Higgs decay into two Z bosons with the subsequent leptonic decay of the Zs provides a clear signature, but its branching ratio below the threshold for the on–shell production of the ZZ pair is low. Nevertheless, since the only irreducible background comes from the ZZ and $Z\gamma$ pair production and the

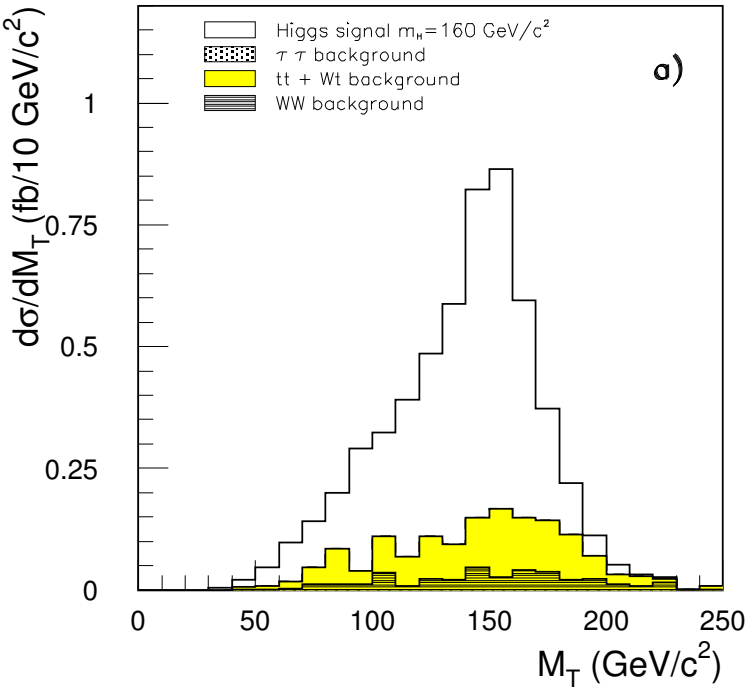

Figure 1.8: *Transverse mass distribution after the kinematical selection of the VBF* $H \to WW \to e\mu + \mathrm{E}_{\mathrm{T}}^{miss}$ *analysis. The units are fb/10 GeV. The Higgs mass is 160 GeV.*

main reducible background comes from $t\bar{t}$ and $Zb\bar{b}$ events where the leptons come from the b–quark decays, requirements of isolation for the leptons in the final state provide a sufficient background rejection. The channel is important up to $m_H \simeq 150$ GeV and above $m_H \simeq 180$ GeV. Between the two values, the opening of the $H \to WW$ into on–shell W bosons causes a drop in the $H \to ZZ$ BR, and thus reduces its possible contribution for the Higgs discovery.

1.4.3 The High Mass Region

The $H \to ZZ \to 4l$ decay provides a powerful, almost background free channel for the Higgs discovery in the mass range between 200 and 600 GeV. Since the Z bosons are boosted, the ZZ and $Z\gamma$ backgrounds can be rejected with an efficient $\mathrm{p_T}$ cut on the reconstructed Z. The signal over background (S/B) ratio is between 5 and 7, with statistical significances $S/\sqrt{B}$ above 20 already with 30 fb^{-1}, if the Higgs mass is below 400 GeV.

The discovery in this channel at high m_H is limited by the decrease of the

inclusive Higgs production cross section. The full coverage of the theoretically allowed m_H range is therefore obtained by looking at more effective channels, like $H \to ZZ \to 2l2\nu$, since the branching ratio of the Z into neutrinos is larger than that into leptons, which allow discovery up to $m_H = 700$ GeV, and the VBF $H \to WW \to l\nu jj$, with a possible discovery up to $m_H = 1$ TeV.

1.4.4 Summary

Figure 1.9 shows the significance obtained with 30 fb^{-1} in the full mass range on the left, without the contribution of the VBF channels in the intermediate mass region [25]. The $H \to ZZ$ channel alone can be used over most of the mass range. As it can be noticed, the most difficult region is that at low masses, which is also that preferred by the electroweak fit of the SM. The plot on the right of the same figure shows only the low and intermediate mass regions with the contributions of the VBF analysis included [27]. The results of the VBF are essential in the low mass region, with the $H \to \tau\tau$ decay, and they increase the total statistical significance in the intermediate mass range with the $H \to WW$ channels.

It should be pointed out that most of the results shown in this chapter have been obtained with a parametrized response of the detector and not with a full simulation. Moreover, the systematics of the QCD background, mainly of the $t\bar{t}$ production, have been estimated with the comparison of parton shower [29], matrix element ([30], [31]) with Next to Leading Order (NLO) generators.

Recently the ATLAS collaboration has performed a more accurate analysis in full simulation, based on an improved detector description, including misalignments, the most recent reconstruction software, and recent Monte Carlo event generators. The results shown in the last chapter of this thesis are part of this new analysis.

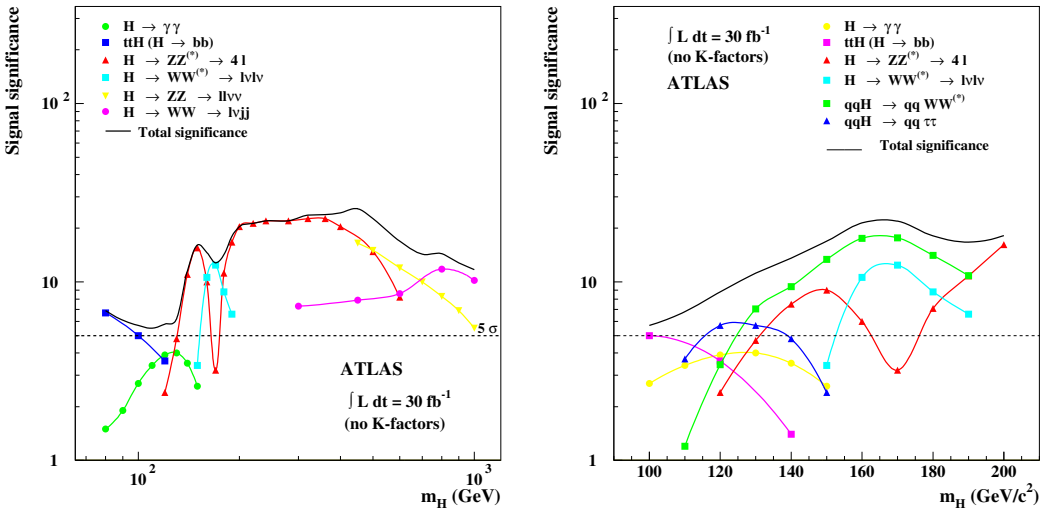

Figure 1.9: *Statistical significance achievable in the different channels discussed in the text as a function of the Higgs mass. On the left: the whole mass range is considered, but the VBF channels are not included. On the right: the low and intermediate mass range, with the vector boson fusion channels included.*

Chapter 2

The ATLAS detector

The stringent requirements for new physics discovery and precise measurements in the SM sector calls for experiments with very good performance. The new CERN accelerator LHC will provide pp collisions at $\sqrt{s} = 14$ TeV *with a design luminosity of* $L = 10^{34}$ cm^{-2}s^{-1}.

In this chapter the ATLAS experimental apparatus is described, which is one of the two general purpose detectors at the LHC.

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [32] is a superconducting proton Synchrotron constructed at CERN inside the 27 Km tunnel used in the past years by LEP. LHC will accelerate from 450 GeV to 7 TeV two beams of protons (or heavy ions) traveling in opposite directions. The existing machines at CERN will provide the first stages of acceleration (Figure 2.1): first, the protons are accelerated up to 50 MeV in the proton LINAC, then the Proton Synchrotron Booster (PSB) boosts them to 1.8 GeV. The Proton Synchrotron (PS) accelerates them up to 25 GeV. Finally, the Super Proton Synchrotron (SPS) will be used to inject 450 GeV protons into the LHC. The LHC dipole magnets house two different magnetic channels in one single twin bore magnet with the same yoke and cryostat. The magnets provide a magnetic field of 8.36 T, which allows the colliding beam particles to reach the design energy of 7 TeV. The design luminosity of 10^{34} cm^{-2}s^{-1} is expected to be reached a few years after the first collisions. After the commissioning of the machine, it is foreseen to have one year at the so called low luminosity ($2 \cdot 10^{33}$ cm^{-2}s^{-1}), which should provide enough statistics to perform the first physics measurements.

The design luminosity for the *pp* collisions will be reached with 2835

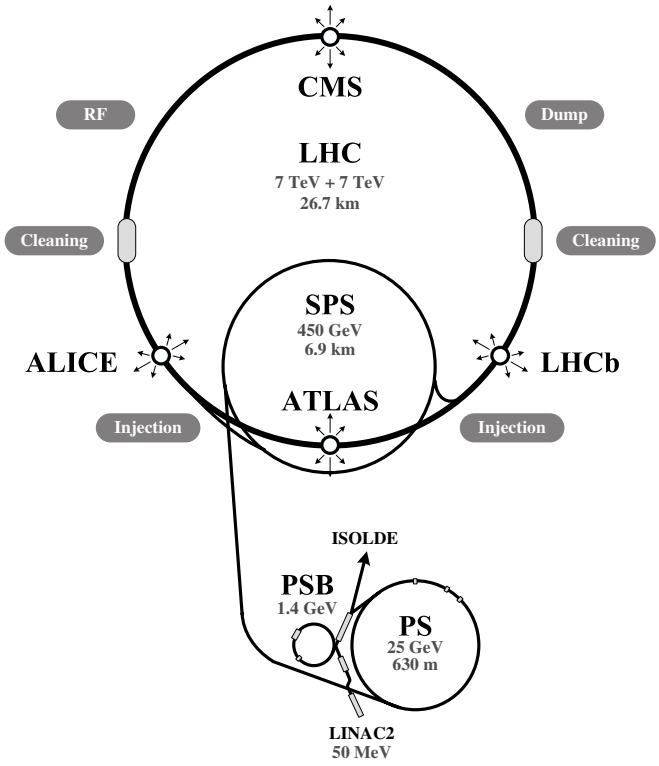

Figure 2.1: *Layout of the LHC.*

bunches crossing at 25 ns intervals, corresponding to a spatial separation of 7.5 m. The number of protons per bunch will be 10^{11}. At the LHC energies, the total inelastic non–diffractive pp cross section is about 70 mb. Since the interesting processes have cross sections that are many orders of magnitude lower (the inclusive Higgs production cross section is well below 1 nb for any Higgs mass), a very selective trigger system is required.

The main parameters of the LHC for pp and heavy–ion collisions are shown in Table 2.1.

Along the perimeter of the LHC, four experiments were built: two general–purpose experiments (ATLAS and CMS [33]), one experiment dedicated to the study of heavy–ion collisions (ALICE [34]), and LHCb [35], dedicated to the study of B mesons decays.

The primary goal of LHC is the study of the origin of the spontaneous symmetry breaking. The discovery (and study) of possible physics beyond the standard model is the other major scientific goal: the possible existence of supersymmetry, fermion compositeness and extra-dimensions can all be investigated with the LHC at the TeV scale.

Parameters	p–p	Pb^{82+}–Pb^{82+}
Beam energy (TeV)	7.0	574
Center of mass energy (TeV)	14	1148
Injection energy (GeV)	450	36900
Bunch spacing (ns)	25	124.75
Particles per bunch	1×10^{11}	7×10^{7}
R.M.S bunch length (m)	0.075	0.075
Number of bunches	2808	592
Initial luminosity ($cm^{-2}s^{-1}$)	10^{33}	1.0×10^{27}
Luminosity ($cm^{-2}s^{-1}$)	10^{34}	1.0×10^{27}
Luminosity lifetime (h)	10	10
Dipole field (T)	8.3	8.3

Table 2.1: *The LHC parameters.*

2.2 The ATLAS Detector Overview

The ATLAS (A Toroidal Lhc ApparatuS) detector [36] is installed at Point 1 of LHC. A picture of the status of the installation on February 2008 is shown in Figure 2.2. The superconducting coils forming the toroidal magnetic system are well visible together with calorimetric system and part of the muon detector. The calorimetry has been designed on the basis of the following criteria:

1. Very good identification and measurement of the four–vectors of electrons and γ's

2. Full hermeticity, both for the EM and hadronic sections

3. Accurate measurement of jets and missing transverse energy.

The coverage in the azimuthal angle ϕ is complete while the geometrical acceptance in pseudorapidity is $|\eta| < 5$. The external muon spectrometer allows high–precision measurements of the muon transverse momentum, even in stand–alone operating mode. This guarantees good p_T measurement also at high luminosity. The tracker has been designed to have high efficiency for non–isolated tracks and to provide electron, photon, τ and b identification.

Figure 2.3 shows an overall schematic view of the detector. The two independent magnetic systems (the solenoidal one in the Inner Detector and the toroidal one in the muon spectrometer) consist of a thin superconducting solenoid surrounding the Inner Detector and eight independent coils arranged with an eight–fold symmetry outside the calorimeters, respectively. The central solenoid provides a 2 Tesla field oriented along the beam axis. The inner

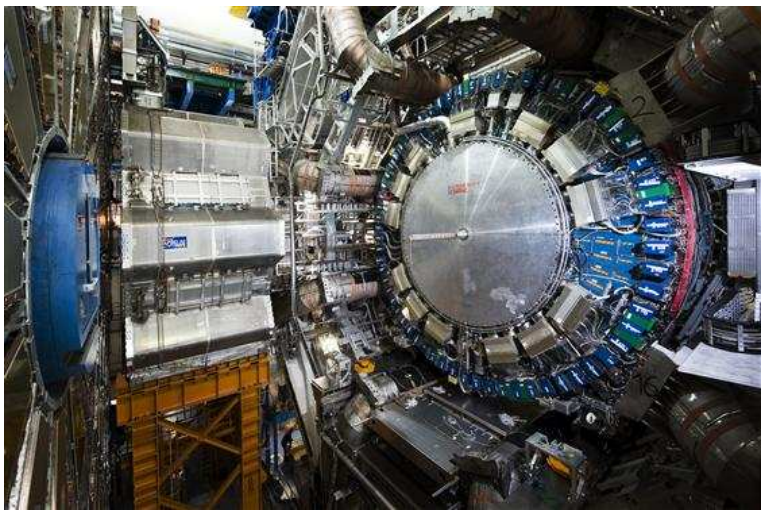

Figure 2.2: *The ATLAS detector as it was on February 2008.*

detector makes use of three different technologies, at different distances from the interaction point. Three inner layers of pixels allow good secondary vertex identification and, together with the four layers of silicon micro-strips, good momentum measurements. The tracking is then completed by continuous straw–tubes detectors with transition radiation detection capability in the outer part.

The calorimetry uses radiation–hard liquid argon (LAr) technology for the EM barrel and end–cap, for the Hadronic End–Cap (HEC) and for the Forward (FCAL) calorimeters. The barrel EM calorimeter is contained in a barrel cryostat, while the EM end–cap, the HEC and the FCAL share the same cryostat in the forward region. In the barrel region the hadronic coverage is provided by the Tile calorimeter (TileCal). Scintillating tiles are used as active material, while the passive material is iron. The TileCal is subdivided in a barrel ($|\eta| < 1$) and an extended barrel ($1 < |\eta| < 1.7$) region. The gap between the two parts is covered by the Intermediate Tile Calorimeter scintillators (ITC), which allow the recover part of the energy lost in the gap.

Outside the calorimeters there is the muon spectrometer. The magnetic field is provided by the 25 m long coils in the central region. The coverage at small angles is completed by two end–cap toroids. The magnetic field bends the particles inside the open structure that constitutes the support for the muon chambers. The multiple scattering is therefore minimized. This allows a very good measurement of the muon momentum with three stations of high precision tracking chambers. The muon detector includes

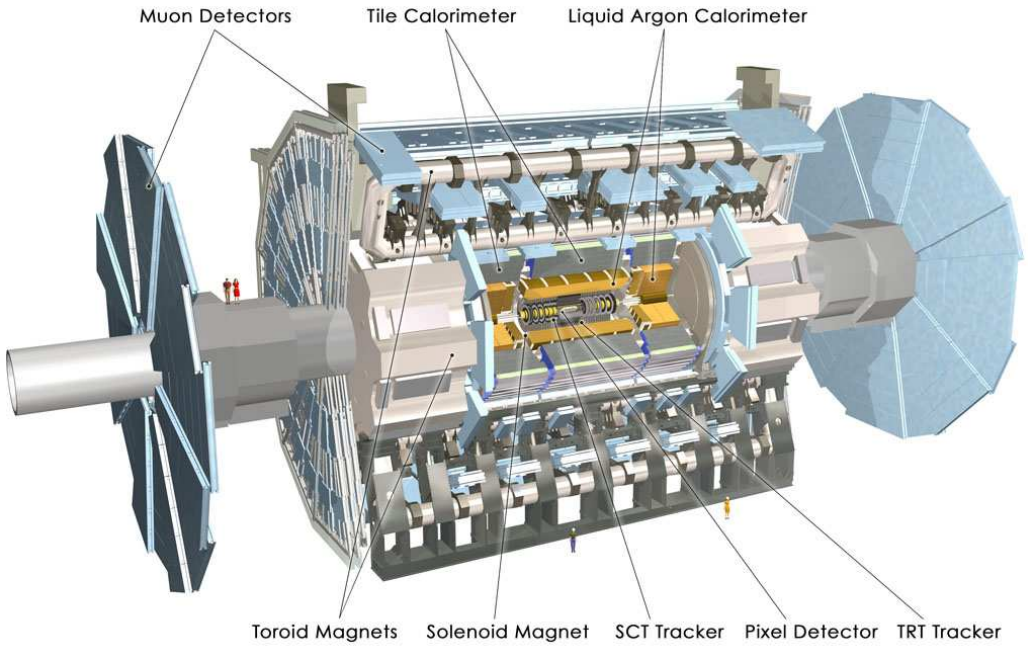

Figure 2.3: *Overall view of the ATLAS detector.*

fast response trigger chambers, which operate in coincidence to provide a fast trigger decision on the muon p_T . The total radius of the ATLAS experiment, from the interaction point to the last muon chamber, is about 11 m. The total longitudinal size is about 46 m, the overall weight is about 7 KTons. Almost 90% of the total ATLAS volume is occupied by the toroids and by the muon spectrometer.

2.3 The Magnetic System

As already mentioned, two different magnetic fields are present within the volume of the detector: the central one, provided by the solenoid, and the outer one, produced by the toroids [37].

The central superconducting solenoid provides a central magnetic field of 2 T, while the peak value at the superconductor face is 2.6 T. In order to obtain the desired calorimetric performances, in particular for photon and electron energy measurements, a careful design to minimize the amount of dead material in front of the calorimeters has been done: the solenoid is placed inside the same vacuum vessel of the LAr calorimeter. The amount of dead material due to the solenoid and the cryostat wall is about one radiation

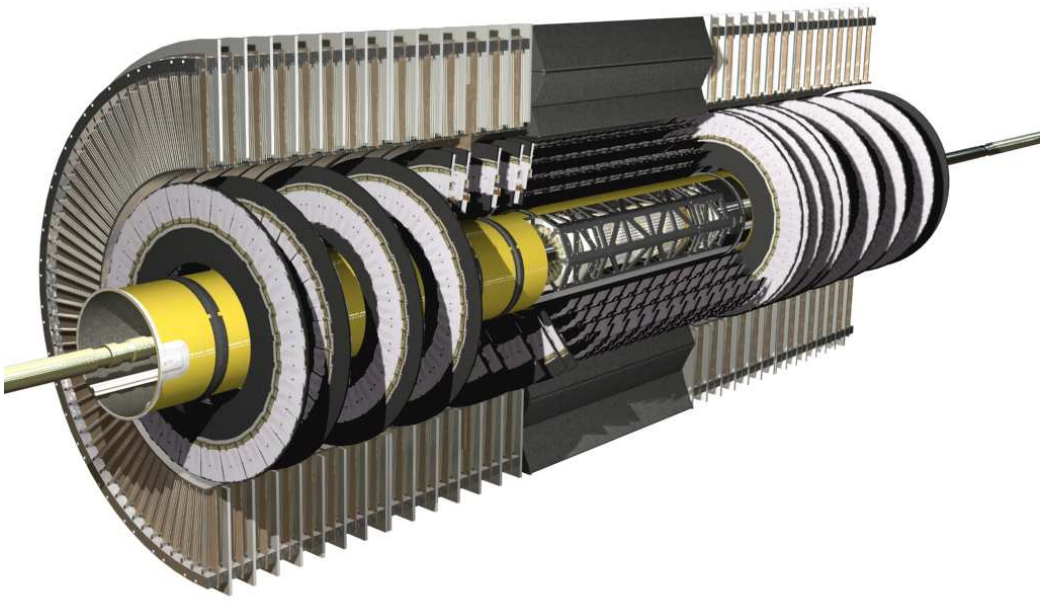

Figure 2.4: *Overall view of the inner detector.*

length.

The magnetic field for the muon spectrometer in the barrel region is provided by a system of 8 coils assembled radially with an eight fold symmetry. The magnetic field in the forward region is obtained with the end–cap coils system, which is rotated by 22.5^o with respect to the barrel coils to provide radial overlap and to optimize the bending power in the interface regions of the two coil systems. The peak magnetic field obtainable in the barrel region is about 4 T. The coils of the barrel are 25 m long and their height is 4.5 m. Each coil is housed in its own cryostat. In the end–cap region there is only one cryostat within which the coils (5 m long and 4.5 m tall) are housed.

2.4 Inner Detector

The strategy used for the ATLAS tracker [38] is to combine few high precision measurements close to the interaction point with a large number of lower precision measurements in the outer radius. The inner detector is embedded in the 2 T magnetic field provided by the central solenoid. The structure of the inner detector is shown in Figure 2.4. Within a radius of 56 cm from the interaction point, pixel and silicon micro-strip technologies offer a fine–granularity, thus a high precision track and vertex reconstruction. A

track typically hits three layers of pixels, which measure both $R-\phi$ and z coordinate, and eight layers of strips (SCT), for a total of 7 tracking points. Two SCT layers form one stereo strip with an angle between them of 40 mrad, allowing the measurement of the three coordinates. In the barrel region, which covers up to $|\eta| = 1$ for a total length of 160 cm, the pixels and SCT are arranged in concentric cylinders around the beam axis, while in the end–cap (up to $|\eta| = 2.5$) they are arranged in disks perpendicular to the beam axis.

A large number of tracking points (36) is provided by the Transition Radiation Tracker (TRT) that also can give e/π separation identifying the transition radiation emitted by electrons. It consists of straw tubes arranged parallel to the beam axis in the barrel region and in wheels around the beam axis in the end–cap. The reduced resolution with respect of the silicon detectors is compensated by the higher radius and by the number of points measured. Therefore, the relative precisions of the measurements of the TRT and pixels/SCT are comparable. The TRT detector is intrinsically radiation hard.

The outer radius of the inner detector cavity is 115 cm, while the total length is 7 m. The layout provides full tracking coverage within $|\eta| < 2.5$, including impact parameter measurement and vertexing for heavy flavors and τ tagging. The expected precision for the detector is [38]

$$\begin{aligned} \sigma_{R-\phi}(\mu\mathrm{m}) &= 13 \oplus \frac{62}{\mathrm{p_T}\sqrt{\sin\theta}} \ , \\ \sigma_{z}(\mu\mathrm{m}) &= 39 \oplus \frac{90}{\mathrm{p_T}\sqrt{\sin\theta}} \ . \end{aligned} \tag{2.1}$$

While the radiation impact is low on the TRT detector, it is not in particular for the pixels, which are more exposed to the radiation since they are closer to the interaction point. The intrinsic radiation weakness of the silicon likely imposes their replacement after a few years of operation, depending on the luminosity profile.

2.5 Calorimeters

The electromagnetic (EM) calorimeter of ATLAS makes use of liquid argon (LAr) as active material and lead as absorber. It covers up to $|\eta| < 3.2$ with a barrel (up to $|\eta| < 1.7$) and an end–cap ($1.5 < |\eta| < 3.2$). In the region at $\eta < 1.8$ (barrel and endcap) there is a presampler detector just behind the cryostat wall. Given the amount of material upstream the EM

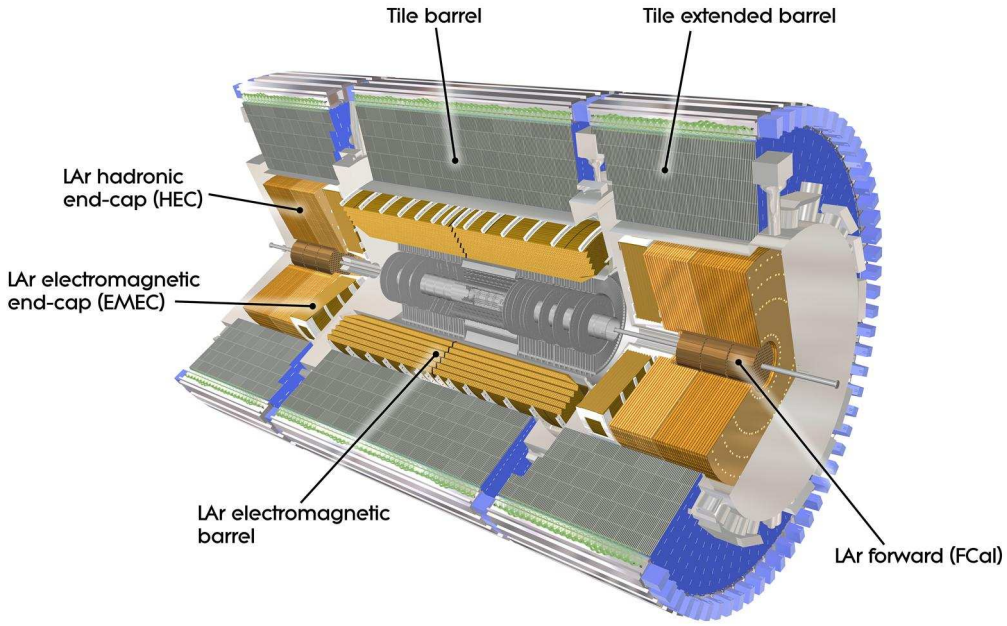

Figure 2.5: *Overall view of the ATLAS calorimeters.*

calorimeter (Inner Detector, cryostat, coil), the presampler is used to correct for the energy loss in it.

The hadronic barrel calorimeter (TileCal) [39] is subdivided into three parts: the central barrel covers up to $|\eta| \simeq 1$, while the two extended barrels cover up to $|\eta| \simeq 1.7$. The energy lost in the gap between the two, where the readout electronics of the EM calorimeter is housed together with the services for the inner part of the detector, can be partially recovered by the gap scintillators (ITC) placed along the internal edge of the extended barrel modules. TileCal uses iron as absorber and scintillating tiles as active material. It is the mechanical support for the LAr cryostats, and it is the return yoke for the central magnetic field flux. The hadronic calorimetry in the forward region uses LAr technology as well. The Hadronic End–Cap (HEC), which covers up to $|\eta| < 3.2$, is placed in the same cryostat of the EM end–cap and of the forward calorimeter (FCAL), which complete the pseudorapidity coverage up to $|\eta| = 5$. The HEC uses copper as absorber, while the FCAL is a dense LAr calorimeter with tubular electrodes in a copper (EM) or tungsten (HAD) matrix.

An overall view of the ATLAS calorimetric system is shown in Figure 2.5, while Table 2.2 shows the details of the segmentation of the calorimeters [40].

EM CALORIMETER	Barrel	End-cap	
Coverage	$\|\eta\| < 1.475$	$1.375 < \|\eta\| < 3.2$	
Long. segmentation	3 samplings	3 samplings	$1.5 < \|\eta\| < 2.5$
		2 samplings	$2.5 < \|\eta\| < 3.2$
Granularity($\Delta\eta \times \Delta\Phi$)			
Sampling 1	0.003×0.1	0.025×0.1	$1.375 < \|\eta\| < 1.5$
		0.003×0.1	$1.5 < \|\eta\| < 1.8$
		0.004×0.1	$1.8 < \|\eta\| < 2.0$
		0.006×0.1	$2.0 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 2	0.025×0.025	0.025×0.025	$1.375 < \|\eta\| < 2.5$
		0.1×0.1	$2.5 < \|\eta\| < 3.2$
Sampling 3	0.05×0.025	0.05×0.025	$1.5 < \|\eta\| < 2.5$
PRESAMPLER	Barrel	End-cap	
Coverage	$\|\eta\| < 1.52$	$1.5 < \|\eta\| < 1.8$	
Long. segmentation	3 samplings	3 samplings	
Granularity($\Delta\eta \times \Delta\Phi$)	0.025×0.1	0.025×0.1	
HADRONIC TILE	Barrel	Extended Barrel	
Coverage	$\|\eta\| < 1.0$	$0.8 < \|\eta\| < 1.7$	
Long. segmentation	3 samplings	3 samplings	
Granularity($\Delta\eta \times \Delta\Phi$)			
Sampling 1 and 2	0.1×0.1	0.1×0.1	
Sampling 3	0.2×0.1	0.2×0.1	
HADRONIC LAr		End-cap	
Coverage		$1.5 < \|\eta\| < 3.2$	
Long. segmentation		4 samplings	
Granularity($\Delta\eta \times \Delta\Phi$)		0.1×0.1	$1.5 < \|\eta\| < 2.5$
		0.2×0.2	$2.5 < \|\eta\| < 3.2$
FCAL CALORIMETER		Forward	
Coverage		$3.1 < \|\eta\| < 4.9$	
Long. segmentation		3 samplings	
Granularity($\Delta\eta \times \Delta\Phi$)		0.2×0.2	

Table 2.2: *Design parameters of the ATLAS calorimeters.*

2.5.1 The LAr Calorimeters

Though there are several differences in the absorber choice and in the read-out, the EM calorimeter, the HEC and the FCAL are grouped in one single subdetector, the LAr calorimeter [41].

The EM Calorimeters

The EM barrel consists of two identical half–barrels separated at $z = 0$ by a 6 mm gap between them, while each of the two EM end–cap is mechanically divided in two coaxial wheels. The outer one covers up to $|\eta| < 2.5$, while

the inner one completes the EM coverage up to $|\eta| = 3.2$. The accordion shape of the electrodes which envelop the lead absorber provides complete ϕ symmetry without azimuthal cracks.

In the barrel region, the lead thickness is optimized as a function of the pseudorapidity to optimize the resolution on the energy measurement. The LAr gap has a constant thickness of 2.1 mm. The geometry of the accordion becomes more complicated in the end–cap, where the amplitude of the accordion waves increase with the radius. Here the absorber have a constant thickness, therefore the size of the LAr gap increases with the radius. Figure 2.6 [41] shows a sketch of the EM structure in the barrel region. As it can be seen, the modules are subdivided in three longitudinal samples, the same in the outer wheel of the end–cap, while the inner-wheel has two samples. The first sample, which is 4.3 radiation lengths (X_0) deep, has a fine segmentation in η, which allows a precise determination of the pseudorapidity of the impinging particle. Since the total amount of material in front of the EM calorimeter at $\eta = 0$ is about 1.7 X_0, there are 6 radiation lengths in front of the second sample, which is the deepest one, with 16 X_0 and a segmentation of $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The third sample has a coarser segmentation in pseudorapidity, and its thickness changes with η. More details about the EM LAr segmentation can be found in Table 2.2.

The signals from the EM calorimeters leave the cryostat through cold–to–warm feedthroughs. They are located between the barrel and extended barrel of the TileCal. The front–end electronics is also located in the gap, where the analog signal (Figure 2.7 [41]) is digitized with fast ADCs. The data are stored in analog pipeline memories, waiting for a possible Level 1 Acceptance (L1A) signal, before digitization. If the event is triggered by the Level 1 (LVL1), the corresponding signals are digitized and 5 samples are extracted and sent to the Read Out Drivers (ROD), where the Optimal Filter (OF) coefficients are applied. Therefore, the energy measurement in each channel is given by

$$E_{cell}(\text{MeV}) = F \sum_{i=1}^{5} a_i (ADC_i - P) \ , \tag{2.2}$$

where F is a conversion factor between ADC counts and MeV obtained with specific calibration runs, P is the pedestal in the cell and a_i are the optimal filtering coefficients. A complete discussion about the optimal filtering can be found in [42]. The linearity of the EM calorimeters has been verified with electrons up to 350 GeV with testbeams (both for the barrel and for the

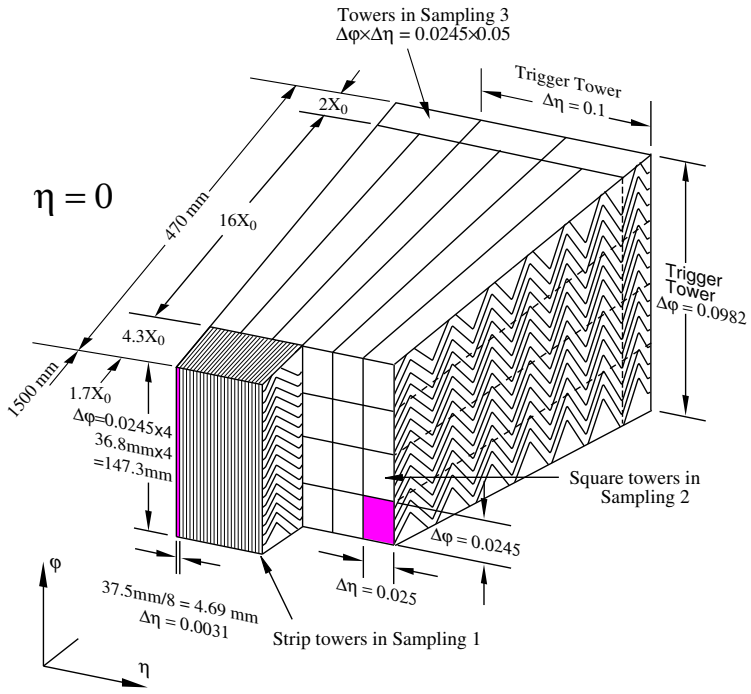

Figure 2.6: *Sketch of the accordion structure of the EM calorimeter.*

extended barrel). The resolution found for the barrel at $\eta = 0.9$ is [41]

$$\frac{\sigma(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus \frac{0.38 \text{ GeV}}{E} \oplus 0.30\% \ , \tag{2.3}$$

which well suites the requirements of the physics.

The Hadronic End–Cap

Each one of the two Hadronic End–Cap consists of two independent wheels of outer radius 2.03 m, for a total coverage of $1.5 < |\eta| < 3.2$. Both wheels consist of an array of copper plates (their thickness is 25 mm in the first wheel, 50 mm in the second). The gap between the plates (8.5 mm) is split by three electrodes into 4 drift spaces of 1.8 mm. The readout electrode is the central one, while the side ones are HV carriers. The scheme is shown in Figure 2.8. The behavior of this configuration is the same of a double gap of 4 mm, but without the drawback associated with large high voltage (typically 2 KV are used instead of 4), nor with large gaps in terms of ion build–up.

Each of the two wheels is composed by 32 identical modules, assembled with fixtures at the peripheric and a central ring. Each wheel is divided into

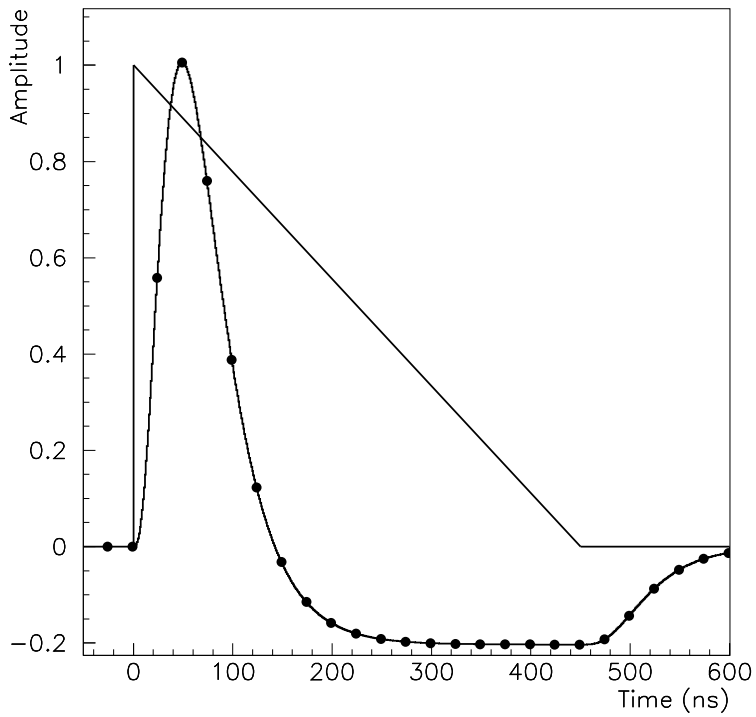

Figure 2.7: *Signal shape of the EM calorimeter as produced in the detector (triangle) and after shaping (curve with dots). The dots represent the position of the successive bunch crossings.*

two longitudinal segments.

Primarily in order to limit the capacitance seen by a single preamplifier, and thus to allow for a fast response, only two gaps are ganged together at the pad level. Miniature coaxial cables running between the sectors carry signals to the preamplifier boards located at the wheel peripheric. Output signals from (typically) four preamplifiers are summed together on the same board. A buffer stage drives the output signal up to the cold-to-warm feedthroughs.

Cells defined in this way are fully projective in azimuth, but only pseudo-projective in η. However, the detector envelope is cylindrical, for sake of mechanical simplicity. To minimize the dip in the material density at the transition between the end-cap and the forward calorimeter (around $|\eta|=$ 3.1), the end-cap EM calorimeter reaches $|\eta|=$ 3.2, thereby overlapping the forward calorimeter.

The HEC standalone resolution for single pions (6–200 GeV) was measured at the testbeam. The results [43] is

$$\frac{\sigma(E)}{E} = \frac{(70.6 \pm 1.5)\%}{\sqrt{E}} \oplus (5.8 \pm 0.2)\% \ . \tag{2.4}$$

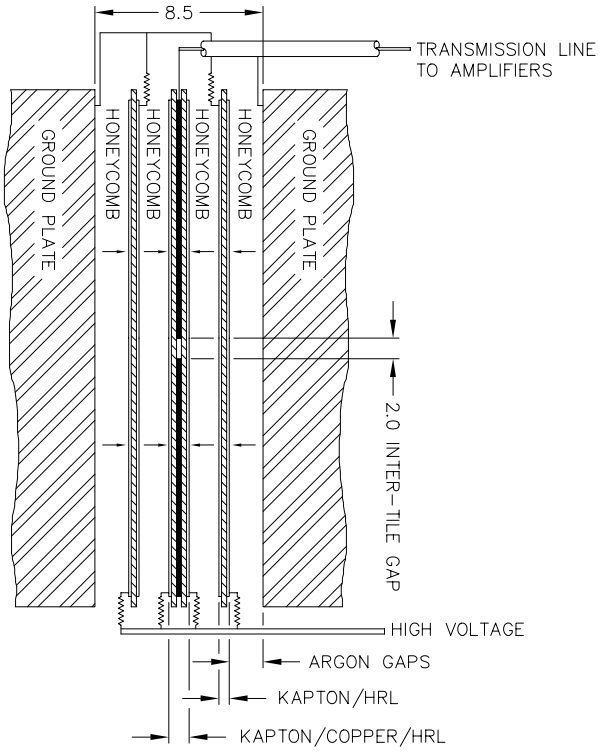

Figure 2.8: *Structure of the readout gap in the Hadronic End–Cap.*

The e/π factor from testbeam is a measurement of the e/h ratio. The result is $e/h = 1.5$.

The LAr Forward Calorimeter

Because of the high level of radiation it has to cope with, the forward calorimeter (FCAL) is built using the intrinsic radiation–hard LAr technology. It is integrated in the forward cryostat together with the EM end–cap and the HEC, and its front face is at about 4.7 m from the interaction point. The presence of this calorimeter provides benefits both in terms of uniformity of the calorimetry coverage and in terms of radiation background for the muon spectrometer.

In order to minimize the amount of neutron albedo in the ID cavity, the front face of the FCAL is recessed by about 1.2 m with respect to the EM calorimeter front face. This severely limits longitudinal space for installing about 9.5 active interaction lengths, and therefore calls for a high-density design, which also avoids energy leakage from the FCAL to its neighbors.

The FCAL consists of three sections: the first one is made of copper,

while the other two are made of tungsten. In each section the calorimeter consists of a metal matrix with regularly spaced longitudinal channels filled with concentric rods and tubes. The rods are at positive high voltage while the tubes and the matrix are grounded. The LAr in the gap is the sensitive medium. This geometry allows for an excellent control of the gaps which are as small as 250 mm in the first section.

2.5.2 The Tile Calorimeter

The central part of the hadronic calorimetry [39] differs from the rest of the calorimetry in that it does not use LAr as active material. Instead, scintillating tiles within an iron structure are used. The structure is periodical along z and the tiles are oriented perpendicular to the beam axis. The tiles thickness is 3 mm while for each period, the total iron thickness is 14 mm. The tiles are read out by two wave length shifting (WLS) fibers, one for each side. The WLS fibers are grouped together in order to reach the desired granularity (three longitudinal samples, with a lateral segmentation of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the first two samples, 0.2×0.1 in the last one) and their signal is read by HAMAMATSU R7877 photomultiplier tubes (PMTs) located on the front–end electronics.

The light emitted inside the tiles in the UV range is shifted to visible violet light by the tiles themselves. The wavelength of the fibers has been chosen so that to maximize the light trapping from the tiles. A light mixer at the end of the fiber bunch makes the readout of the light more uniform on the PMT photocathode. The typical light yield at the PMT photocathode is 0.5 photoelectrons (pe) per tile for a minimum ionizing particle. This corresponds roughly to 20 pe per GeV, which is enough to obtain the desired resolution. A sketch of the optical layout of the TileCal is given in Figure 2.9.

The TileCal is subdivided into one barrel region ($|\eta| < 1$) and two extended barrels ($0.8 < |\eta| < 1.7$, one on each side of the barrel). The gap between them (68 cm) provides space for the services for the inner detector and the front–end electronics of the EM calorimeter. Both the barrel and the two extended barrels are subdivided in 64 modules, one for each ϕ slice ($\Delta\phi \simeq 0.1$). The lateral segmentation of the extended barrel is the same of the barrel, while the longitudinal segmentation differs in the second and third layer.

The gap between the barrel and extended barrel is partially instrumented by the Intermediate Tile Calorimeter (ITC). It is composed of two radial sections attached on the face of the extended barrel. The outer section, 31 cm thick, starts at the outer radius and covers 45 cm in radius. It is followed by the inner section which is 9 cm thick and extends over 45 cm

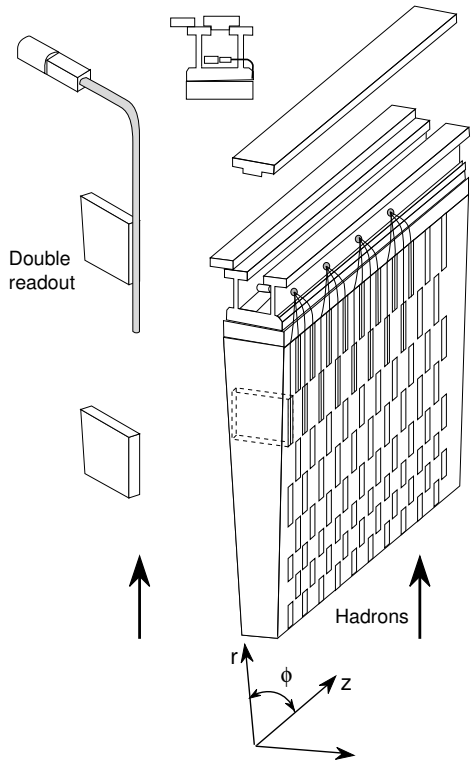

Figure 2.9: *Optical layout of the Tile calorimeter.*

to lower radii. The ITC is extended further inward by a scintillator sheet, covering the inner part of the extended barrel and extending to the region between the LAr barrel and end-cap cryostat over $1.0 < |\eta| < 1.6$. This scintillator samples the energy lost in the cryostat walls and dead material. It is segmented into three sections each covering a range of $\Delta\eta \simeq 0.2$.

The frond–end electronics of one TileCal module is placed on its outer perimeter. For each half–module there is one drawer, which houses the 48 PMTs and the electronics. The signal from the PMT is collected on a 3–in–1 card [44], which shapes the pulse and then sends it to the trigger summation cards and to two different amplifiers with a relative gain factor of 64. After the amplification, the signal is sent to the digitizer boards for digitization. In each drawer there are eight digitizer boards. Each digitizer board provides a double sampling (one for each of the two gains) using two separate 10–bits fast ADCs. The sampled signal is then stored in the pipelines, waiting for the LVL1 decision.

Once the L1A signals is received for one event, the corresponding 7 samplings are extracted from the pipelines and sent to the RODs, where the signal is processed using Optimal Filtering techniques.

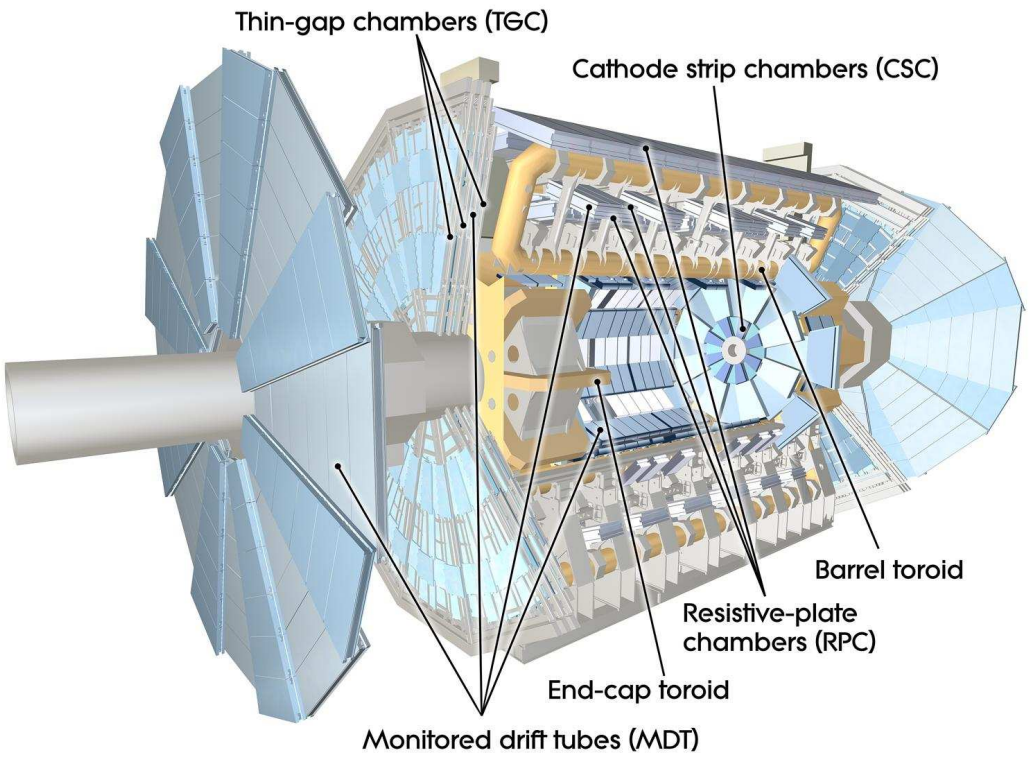

Figure 2.10: *Overall view of the ATLAS muon detector.*

TileCal has three different calibration systems. A hydraulic system can drive a ^{137}Cs source through each tile row, allowing a current measurement of the response of all the optics and PMT chain. A dedicated integrator output on the 3–in–1 card allows to check the calorimeter response independently of the electronic chain. The laser system allows to pulse the calorimeter directly on the PMTs photocathode surface. A system of clear optical fibers different from those used for the read–out of the cells allows splitting the signal between all the PMTs of each module. Finally, the Charge Injection System (CIS) allows to inject a well known amount of charge directly into the 3–in–1 card, and, thus, to check the stability of the electronics chain itself.

Eight modules of the barrel and the extended barrel have been exposed to the SPS beam. The resolution on single pions was found to be [39]

$$\frac{\sigma(E)}{E} = \frac{50\%}{\sqrt{E}} \oplus 1.1\% \tag{2.5}$$

while the e/h factor is 1.36 [39].

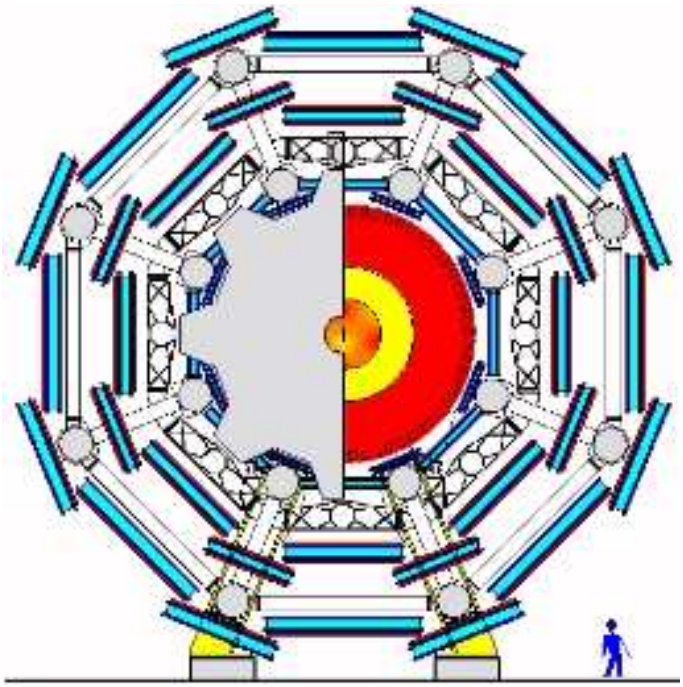

Figure 2.11: *Transverse view of the ATLAS muon spectrometer.*

2.6 Muon Spectrometer

One of the most important features of the muon spectrometer [45] is the possibility of a precise standalone measurement of the muon momentum. The magnetic field provided by the superconducting air–core toroid magnets deflects the muon trajectories that are measured by high precision tracking chambers. The magnetic field in the $|\eta| < 1.0$ range is provided by the barrel toroids, while the region $1.4 < |\eta| < 2.7$ is covered by the end–cap. In the so called transition region ($1.0 < |\eta| < 1.4$) the combined contributions of both the barrel and end–cap provide the magnetic field coverage. The magnetic field is mostly orthogonal to the muon trajectory in the covered pseudorapidity range, while the effect of multiple scattering is minimized.

In the barrel region, the muon chambers are arranged in three cylindrical layers (stations), while in the end–cap they form three vertical walls. The transition region is instrumented with one extra station. Figure 2.10 offers a three dimensional view of the spectrometer. The azimuthal layout follows the magnet structure with 16 sectors (Figure 2.11). The so–called Large Sectors lie between the coils, and they overlap with the Small Sectors, placed in correspondence with the coils themselves.

The choice of the different chambers technology follows the particle flux

Region		station I	station E	station M	station O
Barrel	$\|\eta\| < 1$	MDT		MDT RPC	MDT RPC
End-Caps	$1 < \|\eta\| < 1.4$	MDT TGC	MDT		
	$1.4 < \|\eta\| < 2$	MDT TGC		MDT TGC	
	$2 < \|\eta\| < 2.4$	CSC			MDT
	$2.4 < \|\eta\| < 2.7$	CSC		MDT TGC	

	TRIGGER CHAMBERS		PRECISION CHAMBERS	
Technologies used	RPC	TGC	MDT	CSC
Number of channels	354K	440K	372K	67K
Area (m^2)	3650	2900	5500	27
Time resolution	$< 5\ ns$	$< 7\ ns$	500 ns	<7 ns
Spatial resolution	5-10 mm		80 μm	60 μm

Table 2.3: *Design parameters of the Muon spectrometer.*

expectation in the different regions of the detector. Criteria of rate capability, granularity, aging properties and radiation hardness have been considered. Table 2.3 summarizes the chamber technologies used in the various pseudorapidity regions [45].

The measurement of the track bending coordinate (η) is provided in most of the η region by the Monitored Drift Tubes (MDT), while at large pseudorapidity, the higher granularity Cathode Strip Chambers (CSC) are used. The requirements on the momentum resolution ($\Delta p_T/p_T \simeq 10\%$ at 1 TeV) call for an accuracy of the relative positioning of chambers traversed by a muon track that matches the intrinsic resolution and the mechanical tolerances of the precision chambers. The knowledge of the chamber positioning with an accuracy of 30 μm is required within a projective tower. The accuracy required for the relative positioning of different towers to obtain adequate mass resolutions for multi-muon final states is in the millimeter range. This accuracy can be achieved by the initial positioning and survey of chambers at the installation time. The relative alignment of muon spectrometer, calorimeters and ID will rely on the measurement of the high-momentum muon trajectories.

The MDT chambers are equipped with a in–plane alignment system aiming at a measurement of the tube position displacements, with respect to their nominal positions at the assembly phase, with a precision of better than 10 μm. To achieve this the spectrometer is equipped with a RASNIK system: a laser, mounted at one side of a chamber, projects a pattern to a CCD camera positioned at the other end of the chamber. From the displacement of the pattern–figure respect to what is expected, corrections for chambers deformation can be computed. The chambers for the LVL1 muon trigger system covers the region $|\eta| < 2.4$. Resistive Plate Chambers (RPC) are used in the barrel region, while the Thin Gap Chambers (TGC) are used

in the end–cap. Their first task is to identify without any ambiguity the bunch crossing of the triggered event. This requires a time resolution of better than 25 ns. Next, they have to provide a well defined p_T cutoff for the LVL1 choice. This is obtained considering a window of a size defined by the LVL1 p_T threshold considered on the second RPC (or TGC) station once a super–hit has been obtained in the first station. Finally, the trigger chambers measure the non–bending coordinate (ϕ), in a plane orthogonal to that measured by the precision chambers, with a typical precision of 5–10 mm.

2.7 Trigger and Data Acquisition

The ATLAS trigger and data-acquisition (DAQ) system is based on three levels of on-line event selection. Each trigger level refines the decisions made at the previous level and, where necessary, applies additional selection criteria. Figure 2.12 shows a schematics of the ATLAS trigger system and the expected rate of reduction at each level [46].

Starting from an initial bunch-crossing rate of 40 MHz, the rate of selected events must be reduced to about 100 Hz for permanent storage. The event rate is determined by the total *pp* cross section, and is about 10^9 evt/sec at the nominal luminosity. Hence an overall rejection factor of 10^7 against minimum bias events is required. This strong requirement must match the need of an excellent efficiency for the rare physics processes of interest. Therefore the principal requirement of the Level 1 (LVL1) trigger is that it identifies unambiguously the interesting physics events, while strongly reducing the overall rate.

The LVL1 trigger makes an initial selection based on a reduced granularity information from a subset of detectors. Objects searched by the calorimeter trigger are high p_T electrons and photons, jets, and τs decaying into hadrons, as well as large missing and total transverse energies. High and low transverse momentum muons are identified using only the muon trigger chambers.

The calorimeter selections are based on a reduced-granularity information from all the calorimeters. In the case of the electron/photon and hadron/τ triggers, separated contributions from the EM and hadronic calorimeters are used and then energy isolation cuts can be applied. The missing and total scalar transverse energies used in the LVL1 trigger are calculated by summing over trigger towers. In addition, a trigger on the scalar sum of jet transverse energies is also available.

No tracking information is used at LVL1 due to timing restrictions and the inherently complex nature of the information from the inner detector.

The LVL1 trigger decision is based on a logical combinations of these ob-

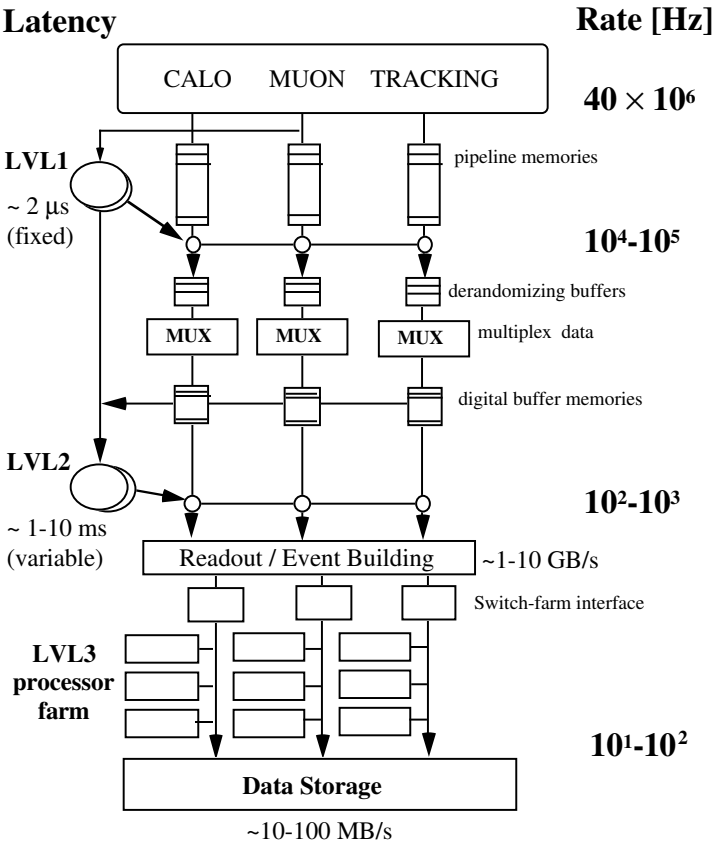

Figure 2.12: *Schematics of the ATLAS trigger system.*

jects. Most of the physics requirements of ATLAS can be met by using fairly simple selection criteria of rather inclusive nature at the LVL1 trigger level. However, the trigger implementation is flexible and it can be programmed to select events using more complicated signatures.

The maximum rate at which the ATLAS front-end systems can accept LVL1 triggers is limited to 75 kHz, upgradeable to 100 kHz. The target rates estimated in trigger performance studies, using trigger menus that meet the needs of the ATLAS physics program, are about a factor of two below this limit.

An essential requirement on the LVL1 trigger is that it should uniquely identify the bunch-crossing of interest. Given the short (25 ns) bunch-crossing interval, this is a non-trivial task. In the case of the muon trigger, the physical size of the muon spectrometer implies times-of-flight comparable to the bunch-crossing period. For the calorimeter trigger, a serious challenge

is that the pulse shape of the calorimeter signals extends over many bunch crossings.

During this time, information for all detector channels are stored in pipeline memories. The LVL1 latency, measured from the time of the *pp* collision until the trigger decision is available to the front-end electronics, is required to be less than 2.5 μs. In order to achieve this, the LVL1 trigger is implemented as a system of purpose-built hardware processors.

Another important functionality of the LVL1 system is the identification of the Regions Of Interest (ROIs) representing the position of the triggered objects in the (η,ϕ) space. This is one of the main peculiarities of the ATLAS trigger system. This information is used to greatly reduce the needed computation time at the LVL2 triggering system and the size of data to be transferred in the system.

The LVL2 runs offline-like algorithms, optimized for the on–line use, using the full granularity information from the inner detector as well as from the muon detectors and calorimetry. However it is structured to process the data belonging only to a spatial window around the ROIs identified by the LVL1 trigger. Simulations showed that this corresponds to roughly 2-5% of the overall ATLAS data size. LVL2 has a maximum latency time of 10 ms, after this time the event is selected and hence moved to the Event Filter system for further processing, or discarded and removed from the Data Flow chain (up to this moment the event fragments have been buffered). The final LVL2 rate is expected to be about 1-2 KHz.

After LVL2, the last stage of the on-line selection is performed by the Event Filter. It employs offline algorithms and methods, slightly adapted to the on–line environment, and uses the most up to date available calibration and alignment information and the magnetic field map. Complete event reconstruction is performed by the Event Filter, which will make the final selection of physics events to be written to mass storage for subsequent off-line analysis. The time available for a decision at the Event Filter is 1 s. The output rate from LVL2 should be reduced by an order of magnitude, giving about 100 Hz. The final event size is expected to be 1 MBytes corresponding to an output data rate of about 100 MByte/s resulting in 10^{15} bytes of data per year.

A more detailed description of the tau-trigger and of the missing transverse energy (E_T^{miss}) trigger, which are fundamental in the analysis presented this thesis, will be given in the next chapter.

2.8 Computing

The complexity and size of the ATLAS experiment impose the use of new paradigms also concerning the data processing once they are made available on mass storage. The event rate of 100 Hz, the size of the events (approximately 1 MB per event), and the number of physicists involved in the analysis require that the data distribution, processing and analysis is carried out according to a multi-tier schema that is well suited to distribute the computing and storage loads among the different participating institutes. Even if similar strategies have been used in the past, it is the first time that this kind of distributed analysis are performed on an ATLAS-size scale requiring the development of completely new software tools [47].

At the output of the event filter, the raw data are transferred to the CERN computing center, known as the Tier-0, the first layer of the ATLAS analysis system. Here a complete copy of the raw data is stored and a first-pass reconstruction is applied producing ESD (Event Summary Data) and AOD (Analysis Object Data). The ESD data-format contains the reconstructed quantities measured by the detector (energy in the calorimeter cells, clusters information, tracks) as well as the reconstructed physics objects (electrons and gammas, jets, taus, muons). The event size foreseen for ESD is about 500 kB.

The small-sized data in AOD format (100 kB per event) are well suited for distribution. Here only the physics objects are recorded. Each event can be characterized by few quantities, like, for example, the number of jets in the event, p_T of the leading jet, lepton multiplicity and so on. This information, produced by Tier-0 and stored in a histogram format in a TAG database, allows a very fast filtering of the datasets.The event size in this case is reduced to about 1 kB.

Tier-0 has also the responsibility to run calibration and alignment algorithms that will be refined in future steps. The distribution of the data to the community is done copying raw data, ESD, AOD and TAG to the Tier-1s. Tier-1s are big regional computer centers spread around the world (at the moment ATLAS foresees 10 of this centers). A copy of the raw data is divided among all the Tier-1s, with each having on average 10% of the entire raw data, while a complete copy of the ESD, AOD and TAG is distributed to each Tier-1. Tier-1s have also the responsibility to reprocess raw data performing more accurate reconstructions. Updated version of ESD, AOD and TAG are therefore constantly produced and spread among the different computer centers.

Most of the physics analysis is performed at the Tier-2 centers: an average of five Tier-2s are directly connected to one Tier-1. These typically receive a

copy of one third of the most updated ESD and AOD data and a complete copy of the TAGs. They have the responsibility for the official Monte Carlo production (the simulated data are stored in the more reliable Tier-1s); the physics groups analysis and the development and refinement of calibration and reconstruction algorithms are also performed at the Tier-2 centers. The physics analysis will be performed mainly on the AOD data sets (with the help of TAGs for pre-selections) or on even more compact Derived Physics Data (DPD), a subsample reduced in size applying stricter event selection, reducing in size the information per object and dropping unwanted data objects.

The multi-tier paradigm is deployed using GRID technology and middleware that completely hides to the physicists the complex multi-tier structure [48].

Chapter 3

Tau and $\mathrm{E}_{\mathrm{T}}^{miss}$ triggers

The selection of events with hadronically decaying tau leptons is challenging due to high QCD background rates at the LHC. On the other hand, efficient selection of events with tau leptons increases the discovery potential of ATLAS in many physics channels, such as the Standard Model or SuperSymmetric Higgs boson among many others. In order to cope with the rate and optimize the efficiency on important physics channels, the tau trigger should be used either with relatively high transverse momentum thresholds alone, or with more relaxed threshold requirements in combination with other triggers, like the missing transverse energy trigger. In this chapter we describe in detail the ATLAS tau trigger and the $\mathrm{E}_{\mathrm{T}}^{miss}$ *trigger to assess their performance. After describing the algorithms of the trigger and the discriminating variables, the final rates and rejection factors are given in the last section.*

3.1 ATLAS Tau Trigger

The identification of the hadronic decays of the tau leptons is based on the selection of narrow isolated jets with low multiplicity in the tracking system. The shower isolation and shape are calculated for both the EM and hadronic calorimeter separately. Given the higher fraction of energy — about 60% — released by the τ-jet in the EM calorimeter and the broader shape of the shower in the hadronic calorimeter, the information obtained from the former is more selective than the one from the latter [49].

In the ATLAS trigger, there are different signatures implemented in the tau trigger, corresponding to different E_{T} thresholds[1]: `tau10i`, `tau15i`,

[1]The first part of the symbol of the signature represents the particle type, the number following is the E_{T} threshold and the letter "i" indicates that an isolation requirement is also applied.

`tau20i`, `tau25i`, `tau35i`, `tau45i`, `tau60`. For `tau10i` and `tau15i` the isolation criteria are only applied at the High Level Trigger. For some E_T thresholds, additional signatures are defined relaxing the isolation criteria (e.g. `tau10`).

3.1.1 Level 1

The L1 tau trigger is a hardware trigger based on the information from electromagnetic (EM) and hadronic calorimeters. It uses trigger towers of an approximate size $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, covering a pseudorapidity range up to $|\eta| < 2.5$, corresponding to the limit of coverage of the Inner Detector and to the high-granularity region of the EM calorimeters.

The algorithm considers a rectangular window of 4×4 towers (0.4×0.4 in $\eta \times \phi$) in both the EM and hadronic calorimeters. For every such window the following quantities are defined, each obtained by summing E_T over a group of towers (see Figure 3.1):

- *core cluster* is the central 2×2 towers region, consisting of both EM and hadronic towers;
- *TauCluster* is the energy in the two most energetic neighboring central towers in the EM calorimeter plus the central 2×2 towers of the hadronic calorimeter;
- *EmIsol* is the energy in the EM isolation ring, i.e. the region surrounding the 2×2 core towers in the EM calorimeter, consisting of 12 EM towers;
- *HadIsol* is the energy in the hadronic isolation ring consisting of 12 hadronic towers behind the EM isolation ring.

The window slides with step of one trigger tower of width of 0.1 units both in the η and ϕ directions.

The L1 object is accepted as a L1 tau trigger candidate if the core cluster is a local E_T maximum, to prevent double counting of clusters by overlapping windows, and if the additional conditions on *TauCluster*, *EmIsol* and *HadIsol* are satisfied [51]. Its position is the center of the 4×4 towers window and its energy is calibrated with a procedure common to jets.

A maximum of eight trigger thresholds are available at L1 for taus. Each threshold is a combination of E_T thresholds for *TauCluster*, *EmIsol* and *HadIsol*.

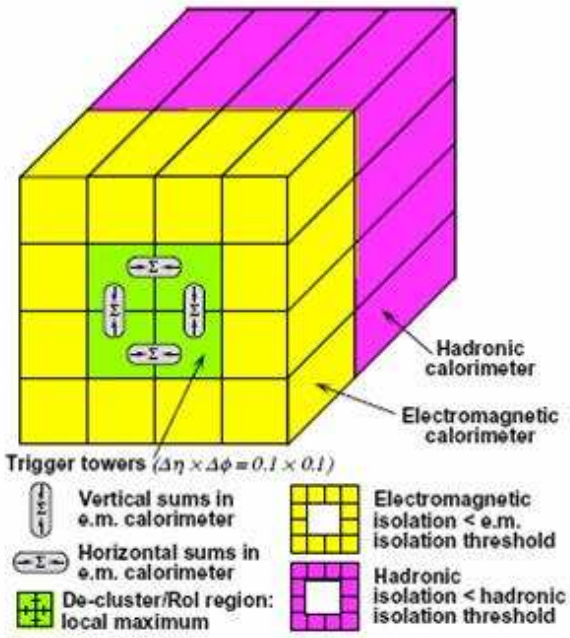

Figure 3.1: *Trigger towers within a tau trigger window.*

3.1.2 Level 2

The level 2 trigger allows the refinement of the L1 signature with full granularity in the calorimeter and the inclusion of tracking information from the Inner Detector. The L2 selection is designed to reject the QCD jet backgrounds by an additional factor of 5-10 exploiting the higher energy resolution and the characteristics of a hadronic tau decay, such as collimation and low track multiplicity [52].

Level 2 Calorimeter

The L2 tau trigger makes use of calorimeter shower shape variables and the cluster energy to select tau candidates. Shape variables are calculated using only the second EM sample layer while the cluster energy is calculated with all available EM and hadronic layers. The second layer is used because it has the largest depth in the radial direction and therefore provides a better shower characterization since each cell has a granularity 0.025×0.025 in $\Delta\eta \times \Delta\phi$.

Three different rectangular window sizes are defined, centered on a *seed* cluster, with areas of $\eta \times \phi = 0.1 \times 0.1$ (narrow, *Nar*), 0.2×0.2 (wide, *Wid*), and 0.3×0.3 (normal, *Nor*).

The algorithm consists of several steps. First of all, the *seed* of the cluster

is found by refining the L1 Region of Interest (RoI) position using the second EM sample. The algorithm looks at cells in the *Nor* region around the L1 RoI position, and finds the cell with the highest energy deposition. In a *Wid* region around the most energetic cell, the cluster position (*seed*) is defined as the energy weighted mean position of the cells in this window. Then, shape variables are reconstructed using different window sizes around the seed, as described in the following. At the same time, the total energy in all calorimeter samples is computed. Finally, the total energy is corrected with a simple sampling calibration.

The rejection variables used by the L2 Calorimeter are the following:

- *EMRadius* is the energy weighted squared radius of the cluster, which is obtained from the sum of the individual cell energies weighted with the square of the cell distance from the *seed*. It is calculated in a *Nor* region around the *seed*, in the second sample of the EM calorimeter, i.e.:

$$EMRadius = \frac{\sum_{Nor} \mathrm{E}_{cell} \cdot R_{cell}^2}{\sum_{Nor} \mathrm{E}_{cell}} \ . \tag{3.1}$$

- *IsoFrac* is the difference in energy between the *Nar* and *Wid* region, normalized to the *Wid* region. It is calculated in the second sample of the EM calorimeter. The definition is:

$$isoFrac = \frac{\sum_{Wid} \mathrm{E}_{cell} - \sum_{Nar} \mathrm{E}_{cell}}{\sum_{Wid} \mathrm{E}_{cell}} \ . \tag{3.2}$$

- *StripWidth* is the width of the energy deposition, defined as the energy weighted standard deviation in η. It is calculated in a *Nor* region around the seed, in the second sample of the EM calorimeter[2]. The formula is:

$$stripWidth = \sqrt{\frac{\sum_{Nor} \eta_{cell}^2 \cdot \mathrm{E}_{cell}}{\sum_{Nor} \mathrm{E}_{cell}} - \left[\frac{\sum_{Nor} \eta_{cell} \cdot \mathrm{E}_{cell}}{\sum_{Nor} \mathrm{E}_{cell}}\right]^2} \ . \tag{3.3}$$

[2]The name of this variable comes from the offline quantity that is calculated using the first, fine segmented, sample of the EM calorimeter. At L2 the name of this variable is kept.

- *EtCalib* is the calibrated total transverse energy, calculated in the EM and hadronic calorimeters, in a *Nor* region around the seed.

The distributions of *EMRadius*, *IsoFrac* and *stripWidth* are shown in Figure 3.2 [52] for signal taus and background jets passing low E_T and high E_T L1 tau selections.

Level 2 Tracking

The standard L2 track reconstruction is used at L2 in the tau slice. No track reconstruction is attempted in the Transition Radiation Tracker. Tracks with $p_T > 1.5$ GeV are reconstructed in a rectangular region $\Delta\eta \times \Delta\phi = 0.6 \times 0.6$ centered on the L2 Calorimeter seed.

The output of the L2 Tracking algorithm is a list of tracks found in the RoI. Two window sizes are defined, corresponding to a distance $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ of 0.15 (*Core*) and 0.3 (*Nor*) with respect to the highest p_T track. An isolation ring (*Iso*) with ΔR between 0.15 and 0.3 is also defined.

The following selection criteria are then calculated from the track list:

- *Pt leading* is the p_T of the highest p_T track found. By requiring a minimum p_T this criterion also effectively requires that at least one track is found in the RoI.

- *Pt Iso/Core* is the ratio of the scalar sum of track p_T in the *Core* and *Iso* region: $\sum p_T^{iso} / \sum p_T^{core}$.

- *N Slow tracks* is the number of slow tracks found in the *Core* region. A slow track is defined as a track with p_T below some threshold, typically 7.5 GeV. The rejection power of this variable might depend on the pileup conditions.

- *Charge* is defined as the absolute value of the sum of charges of all tracks found in the *Nor* region.

- *N Tracks* is the total number of tracks found in the *Nor* region.

The L2 track cuts further enrich the tau sample requiring at least one track to be found with a minimum *Pt leading*, a maximum amount of energy deposited in the *Iso* region, a maximum number of slow tracks. The requirements on charge and total tracks are set very loose, partly because of the higher amount of fake tracks present at L2 due to a less refined reconstruction with respect to the offline and partly because for single tau final states (e.g. $W \to \tau\nu$) an unbiased track distribution is needed to extract the number of signal events.

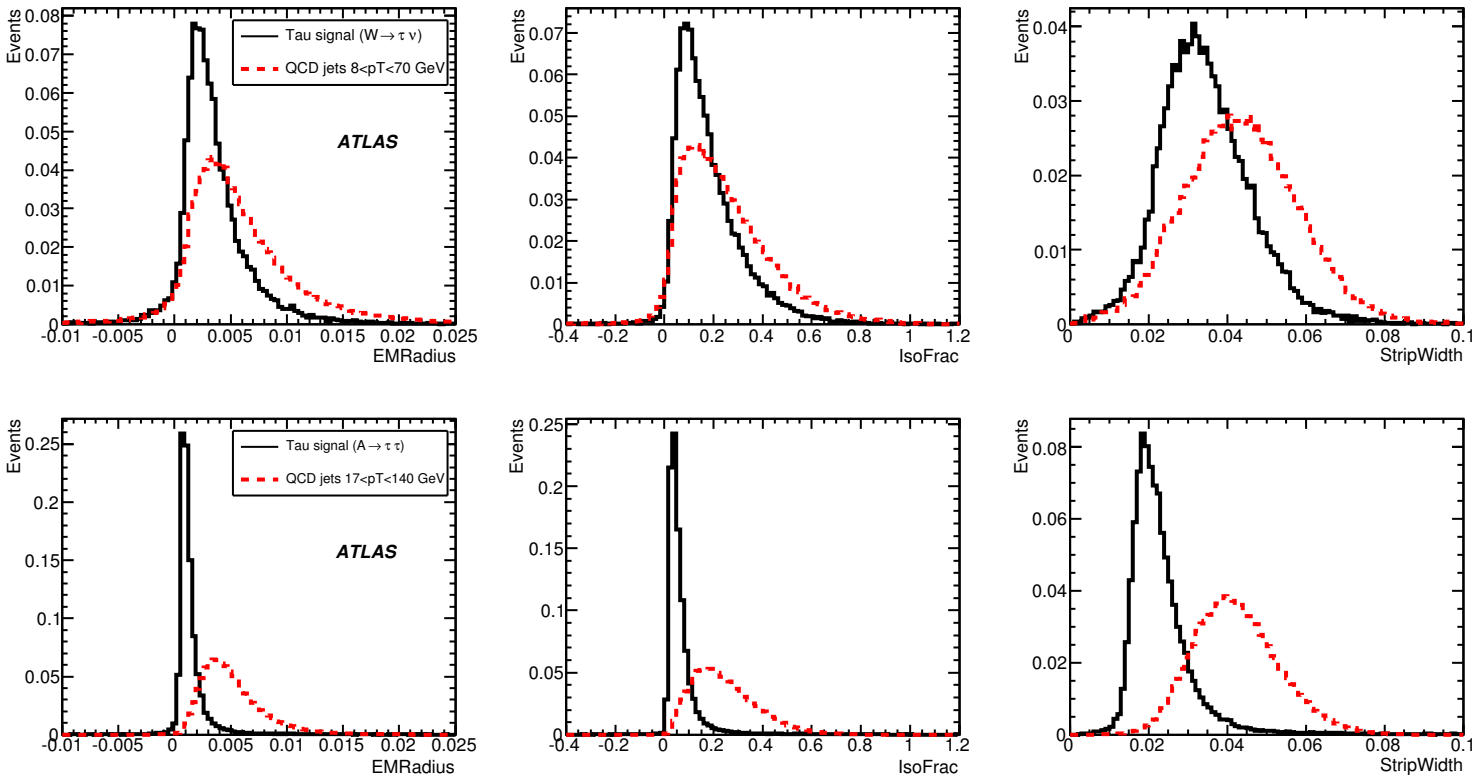

Figure 3.2: *Calorimeter L2 variables from L1 RoI to distinguish QCD background from low* $\mathrm{E_T}$ *taus (top) and from high* $\mathrm{E_T}$ *taus (bottom) using* $W \to \tau\nu$ *and* $A \to \tau\tau$ *(with* $m_A = 800\,GeV$*) as signal.*

The distributions of *Pt leading*, *N Slow tracks* and *Pt Iso/Core* can be seen in Figure 3.3 [52] for signal taus and background jets passing low E_T and high E_T L1 and L2 calorimeter tau selections.

3.1.3 Event Filter

At the Event Filter (EF), as in L2, the selection starts with a calorimeter algorithm first, and secondly with a tracking algorithm.

The EF calorimeter algorithm collects cells in a rectangular region $\Delta\eta \times \Delta\phi = 0.6 \times 0.6$ centered around the L2 seed, it recalculates the seed direction, and cells in a cone around that seed are used to reconstruct energy and calorimeter identification variables. The variables of the offline calorimeter-based identification algorithm [53] are used. Tracking is performed in a rectangular region $\eta \times \phi = 0.4 \times 0.4$ centered around the L2 seed.

Finally, an overall calibration is applied to all cells, and after that a tau specific jet calibration is applied to the tau object. This final calibration is derived from simulated samples of high and low p_T tau lepton sources, like Z and Higgs bosons.

The following variables are currently used for tau identification at EF:

- *EMRadius* is an energy weighted radius, which exploits the small transverse shower profile of tau leptons. The mathematical expression used is

$$EMRadius = \frac{\sum_i E_{T,cell} \cdot \Delta R_{cell}}{\sum_{Nor} E_{T,cell}} \ . \tag{3.4}$$

  where sum extends to all layers of the EM calorimeter up to the second included.

- *IsoFrac* is the energy deposited in an annular region between $0.1 < \Delta R < 0.2$ divided by the energy deposited in $\Delta R < 0.3$. A similar expression is given in Eq. 3.2, only the square regions should be replaced by the annuli in this case.

- *N Tracks* is the number of associated tracks with $p_T > 2\,\text{GeV}$ found in a $\Delta R < 0.2$ region around the seed.

- *Pt leading track* is the highest p_T track among the tracks found in a $\Delta R < 0.2$ region around the seed.

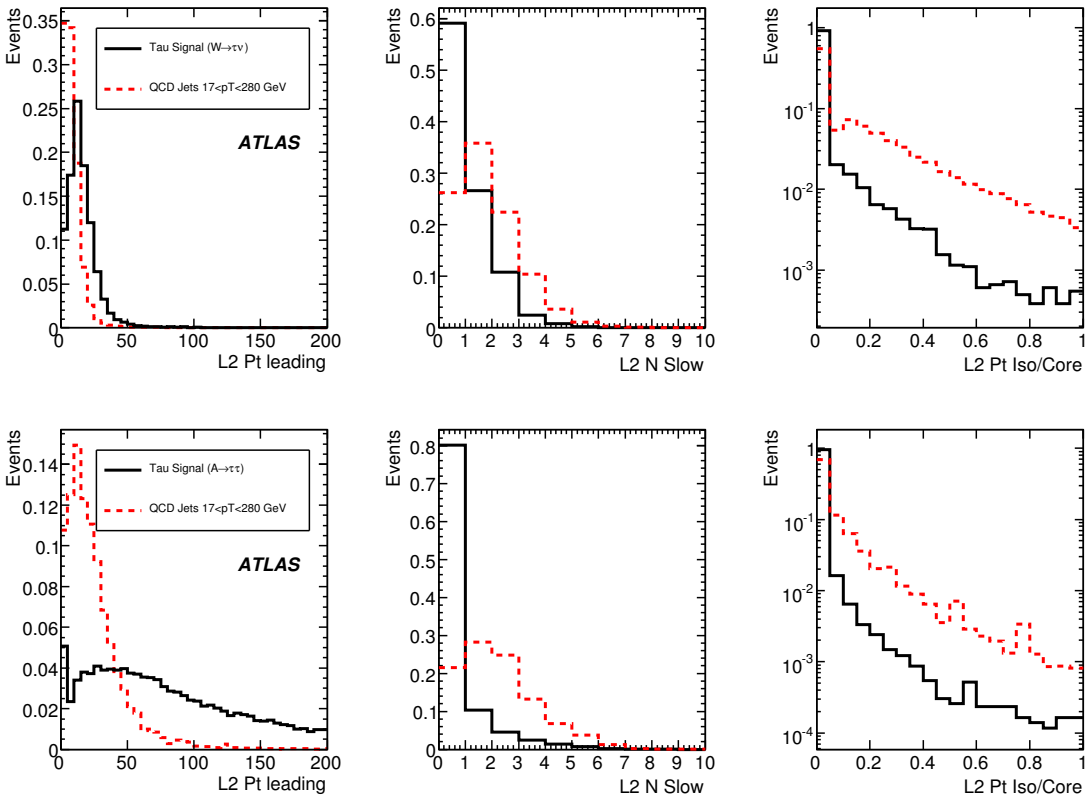

Figure 3.3: *Tracking L2 variables to distinguish QCD background from low* $\mathrm{E_T}$ *taus (top) and from high* $\mathrm{E_T}$ *taus (bottom) using* $W \to \tau\nu$ *and* $A \to \tau\tau$ *(with* $m_A = 800\,GeV$*) as signal.*

Rapidity coverage	Mean value of E_T^{miss}
All	0.9 GeV
$\|\eta\| < 5$	5.6 GeV
$\|\eta\| < 4$	8.8 GeV
$\|\eta\| < 3$	12.5 GeV

Table 3.1: *Mean value of* E_T^{miss} *as a function of rapidity coverage.*

- *EtCalib* is the energy calculated in all EM and hadronic cells found in the $\Delta R < 0.3$ region around the seed, and calibrated with the procedure described in [54] and an additional tau specific jet calibration.

The variables *EtCalib*, *IsoFrac* and *EMRadius* and are shown in Figure 3.4 [52] for low and high E_T taus. No noise suppression on cells ($E > 2\sigma(noise)$) is applied by the trigger, as opposed to what happens in the offline algorithms. The addition of electronic noise in the calorimeter cells gives a change in the calorimeter identification variables which is generally small.

In the EF, a simple cut based selection is applied on the variables described above.

3.2 ATLAS E_T^{miss} Trigger

3.2.1 Level 1

In ATLAS the missing transverse energy (E_T^{miss}) trigger covers the region $|\eta| < 4.9$ [55], corresponding to the limit of the Forward Calorimeter.

E_T^{miss} is a global variable of the event. Crucial for the E_T^{miss} trigger performance is the rapidity coverage of the detector. In Table 3.1 the mean value of E_T^{miss} in QCD events is given as a function of η.

The symbol for the signature is `xe` and the number that follows represents the E_T threshold.

The basic unit of L1 trigger algorithms is the *jet element*, a reduced granularity object formed by summing over 2×2 trigger towers to obtain a 0.2×0.2 granularity in $\Delta\eta \times \Delta\phi$. Each jet element is multiplied by an appropriate geometrical constant to obtain its transverse-energy components E_X, E_Y and ΣE_T. Then the algebraic sum is performed in order to get the total E_X and E_Y and a lookup table is used to perform the final quadrature sum of E_X and E_Y.

Muons are not taken into account in L1 estimate of E_T^{miss}.

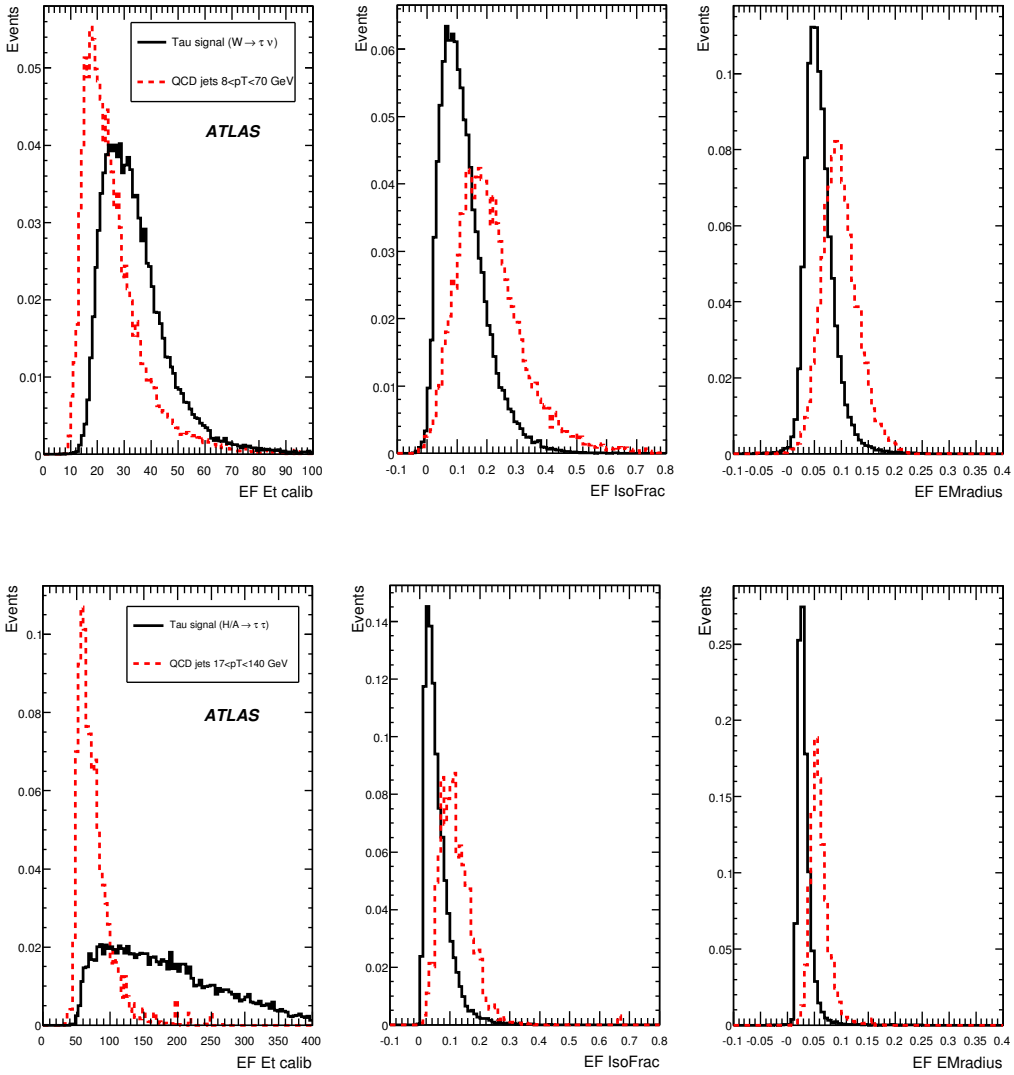

Figure 3.4: *Calorimeter EF variables to distinguish QCD background from low* p_T *taus (top) and from high* p_T *taus (bottom) using* $W \to \tau\nu$ *and* $A \to \tau\tau$ *(with* $m_A = 800\,GeV$*) as signal.*

The E_T^{miss} turn-on curve is largely limited by L1 resolution. Figure 3.5 [56] shows turn-on curves of the L1 E_T^{miss} trigger with respect to the reconstructed values for leptonic $t\bar{t}$ events. Vertical offsets are seen for trigger thresholds of 50 GeV and below because a considerable number of events pass the trigger thresholds for reconstructed E_T^{miss} values of 0 GeV. At low E_T^{miss} , the L1 trigger overestimates the reconstructed E_T^{miss} values. The shape and offset of the turn-on curve for L1_XE100 are reasonable, but the 20% efficiency at 100 GeV means that the reconstructed energy value is now underestimated by the trigger system for large E_T^{miss} .

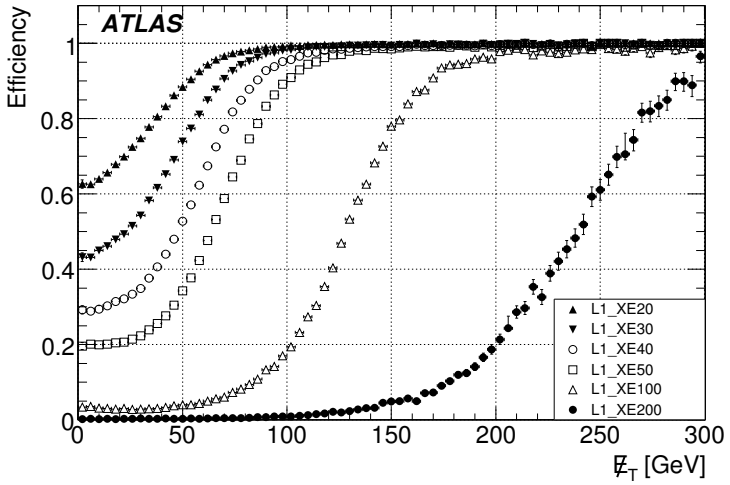

Figure 3.5: *Turn-on curves for* E_T^{miss} *triggers at L1, with respect to reconstructed values, for the leptonic* $t\bar{t}$ *samples.*

3.2.2 Level 2 and Event Filter

For both L2 and EF a final strategy has not yet been established [57]. At the second trigger level all the information is available but limited to the L1 RoI.

At the present the implementation of the L2 algorithm is achieved by using the calorimeter information in the Regions-of-Interest provided by L1 trigger. For a correct estimate of E_T^{miss}, the muon energy has to be included.

The default EF algorithm sums all calorimeter cells and applies a basic hadronic calibration multiplying E_T^{miss} and ΣE_T by a constant related to the hadronic/electromagnetic calorimeter energy fraction in a jet. The information on muons has to be added to calculate the EF E_T^{miss}.

The timing requirements, together with the rejection power of the HLT algorithms, are crucial for the final implementation choice.

3.2.3 E_T^{miss} and ΣE_T Resolution

In Figure 3.6 [57] the E_T^{miss} and ΣE_T resolution, defined from the width of $E_T^{miss} - E_T^{miss,True}$ in bins of ΣE_T, is shown as a function of true ΣE_T in $t\bar{t}$ events for L1, EF and offline algorithms. Since the L2 algorithm takes L1 results and apply relatively small corrections, the L2 resolution is about the same as for L1 and is therefore not shown. For a true ΣE_T value of 500 GeV, the L1 resolution on E_T^{miss} measurements is about 25 GeV, which is a factor two larger than that one achieved offline. The EF resolution, for both E_T^{miss} and ΣE_T lies in between L1 and offline resolution values.

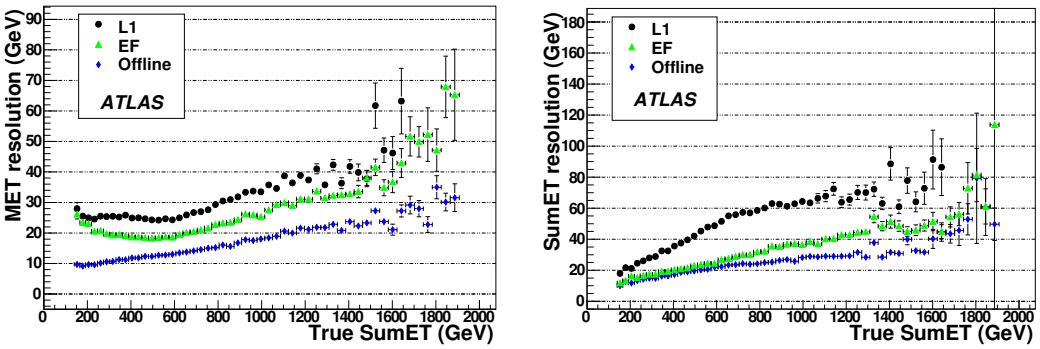

Figure 3.6: *L1, EF and offline resolution for* E_T^{miss} *(left) and* ΣE_T *(right) as a function of true* ΣE_T *in* $t\bar{t}$ *events.*

3.3 Trigger Performances

The tau trigger efficiency is optimized with respect to generated tau leptons decaying hadronically, where the visible tau momentum ($p^{\alpha}_{\mathrm{vis}} = p^{\alpha}_{\tau} - p^{\alpha}_{\nu}$) is in the sensitive region of the detector ($|\eta| \leq 2.5$) and is also greater than the nominal E_T threshold requirement for a given signature.

The tau trigger is further optimized to select those tau leptons which are likely to be reconstructed with the offline tau algorithms. In the signal efficiency calculation an event that fires the τ trigger is considered only if that τ is also selected by the offline tau reconstruction algorithms.

Using $W \rightarrow \tau\nu$ and $A(800) \rightarrow \tau\tau$ samples, the efficiency is plotted as a function of the Monte Carlo visible E_T for different tau trigger signatures in Figure 3.7 for the L1, in Figure 3.8 for the L2, and in Figure 3.9 for the

EF [52]. Figure 3.10 shows the overall efficiency for different tau trigger signatures.

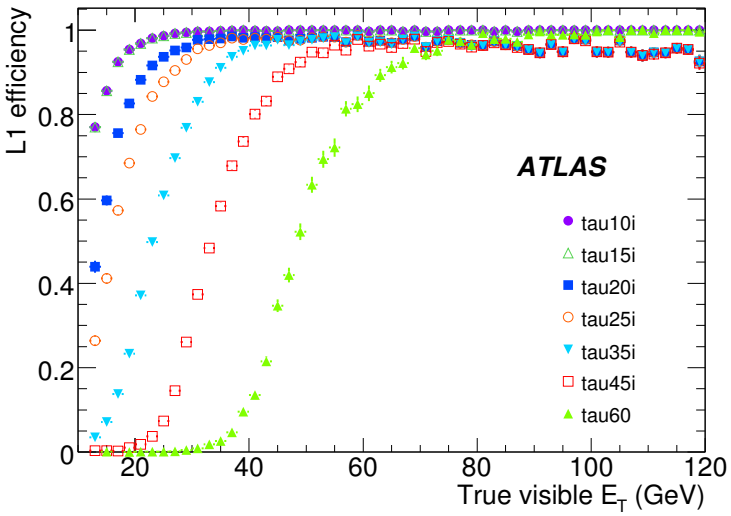

Figure 3.7: *Efficiency curves for different L1 tau signatures. Efficiency is given per tau.*

Tau triggers can be combined with other signatures. ATLAS trigger menu includes tau+ E_T^{miss} triggers. This group of triggers covers a wide spectrum of physics channels. At low luminosity, when the trigger rejection can be relaxed, the selection of events with $W \rightarrow \tau\nu$ is the priority. $t\bar{t}$ events with τs in final states are also selected by this trigger. Such events are characterized by relatively soft E_T range of taus as well as low E_T^{miss}. The tau+E_T^{miss} signatures at designed luminosity are optimized for Higgs (neutral or charged) searches as well as for searches of new exotic particles or processes. For items `tau+xe`, the true E_T^{miss} as well as the reconstructed offline E_T^{miss} are required to be above the chosen trigger threshold.

ATLAS will collect data at different luminosities, starting from $10^{31}\text{cm}^{-2}\text{s}^{-1}$ up to $10^{33}\text{cm}^{-2}\text{s}^{-1}$ within several months. While the input rate rapidly increases during the first period, the maximal output rate of the complete trigger is fixed at 200 Hz, subject to storage capacities and recording speed. In the Table 3.2 some of tau and tau+E_T^{miss} trigger efficiencies for $100\,pb^{-1}$ for various physics signals are summarized, and the corresponding rates estimated on minimum bias samples are shown in the Table 3.3 [52].

For the Higgs search in the H$\rightarrow \tau\tau \rightarrow hh$ channel studied in this thesis the `tau35i_xe40` trigger has been chosen for the analysis. This trigger is foreseen to run unprescaled during the data taking period at luminosity $\mathcal{L} =$

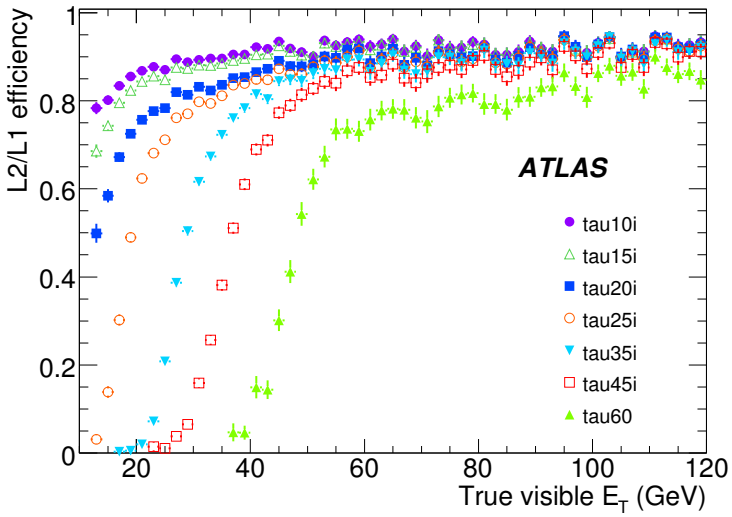

Figure 3.8: *Efficiency curves for different L2 Calorimeter tau signatures. Efficiency is given per RoI passing L1 selection.*

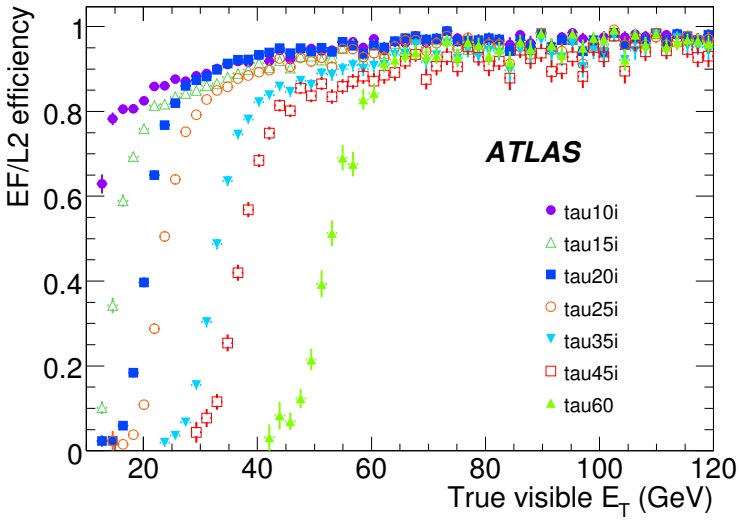

Figure 3.9: *Efficiency curves for different EF tau signatures. Efficiency is given per RoI passing L1 and L2 selection.*

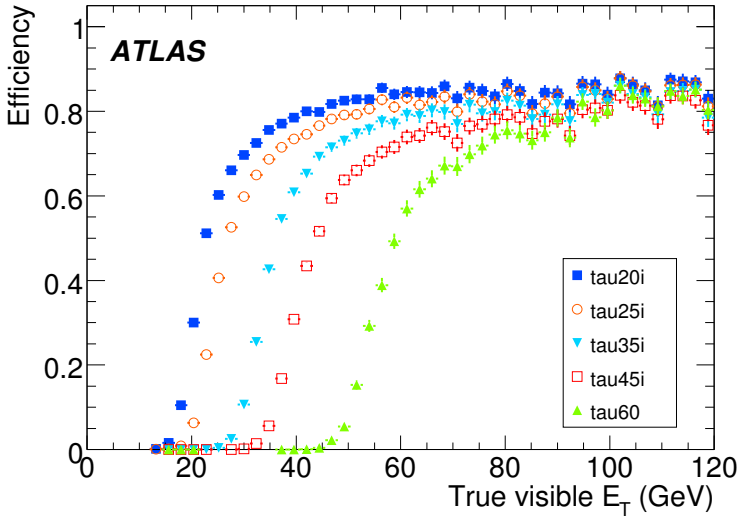

Figure 3.10: *Overall trigger efficiency (L1 + L2 + EF) for different tau signatures.*

Trigger Item	$W_{\tau\to hX}$	$Z_{\tau\tau}$	$t\bar{t}$	$A_{\tau\tau}(800)$	$H_{\tau\tau\to\ell hX}$	$H_{\tau\tau\to hhX}$
tau10i	78.7 ± 0.4	89.8 ± 0.2	91.1 ± 0.4	96.4 ± 0.2	92.1 ± 0.3	95.8 ± 0.1
tau15i	74.1 ± 0.4	86.0 ± 0.3	89.1 ± 0.5	96.1 ± 0.2	90.0 ± 0.3	94.1 ± 0.1
tau20i	68.5 ± 0.5	79.9 ± 0.3	83.2 ± 0.6	89.8 ± 0.2	85.0 ± 0.4	89.8 ± 0.2
tau25i	66.5 ± 0.6	76.0 ± 0.4	80.1 ± 0.7	89.0 ± 0.3	82.0 ± 0.4	87.1 ± 0.2
tau35i	65.8 ± 0.9	70.0 ± 0.6	77.4 ± 0.9	87.5 ± 0.3	78.2 ± 0.5	82.0 ± 0.2
tau45i	72.1 ± 1.5	68.5 ± 0.9	76.0 ± 1.2	86.1 ± 0.3	75.8 ± 0.7	78.5 ± 0.3
tau60	77.5 ± 2.6	74.4 ± 1.5	74.7 ± 1.7	91.4 ± 0.2	76.1 ± 0.9	77.5 ± 0.4
tau15i_xe20	56.3 ± 0.6	48.4 ± 0.8	80.1 ± 0.7	92.7 ± 0.2	80.8 ± 0.5	80.2 ± 0.3
tau20i_xe30	45.4 ± 0.9	38.6 ± 1.2	70.1 ± 0.9	84.5 ± 0.3	73.9 ± 0.6	73.2 ± 0.4
tau25i_xe30	44.2 ± 1.0	38.0 ± 1.3	67.5 ± 1.0	83.8 ± 0.3	71.4 ± 0.7	71.1 ± 0.4
tau35i_xe20	55.1 ± 1.2	42.0 ± 1.2	68.7 ± 1.1	84.5 ± 0.3	70.8 ± 0.7	69.5 ± 0.4
tau35i_xe30	47.7 ± 1.5	38.5 ± 1.7	63.3 ± 1.2	82.3 ± 0.3	69.2 ± 0.8	66.9 ± 0.4
tau35i_xe40	42.1 ± 2.5	39.2 ± 2.4	58.0 ± 1.4	80.8 ± 0.4	68.3 ± 1.0	65.3 ± 0.5
tau45i_xe20	60.1 ± 2.1	43.8 ± 1.7	67.3 ± 1.4	83.2 ± 0.3	69.0 ± 0.9	66.7 ± 0.4

Table 3.2: *Unprescaled signal efficiency. The parameter m_H is taken equal to 120 GeV in the last two columns. The quoted errors are only statistical.*

Trigger	Level 1		Level 2		Event Filter	
Selection	Events	Rate (Hz)	Events	Rate (Hz)	Events	Rate (Hz)
tau10i	48854	14290	14066	4114	5413	1583 ± 21
tau15i	48854	14290	9764	2856	2953	864 ± 16
tau20i	16298	4767	3322	972	1229	359 ± 10
tau25i	9404	2751	1565	458	631	185 ± 7
tau35i	3149	921	473	138	215	63 ± 4
tau45i	750	219	154	45	79	23 ± 3
tau60	262	77	35	10.2	25	7.3 ± 1.5
tau15i_xe20	5625	1645	1329	389	198	58 ± 4
tau20i_xe30	525	154	139	40.7	23	6.7 ± 1.4
tau25i_xe30	385	113	83	24.3	17	5.0 ± 1.2
tau35i_xe20	784	229	129	37.7	42	12.3 ± 1.9
tau35i_xe30	203	59	41	12.0	10	2.9 ± 0.9
tau35i_xe40	58	17	15	4.4	3	0.9 ± 0.5
tau45i_xe20	278	81	54	15.8	19	5.6 ± 1.3

Table 3.3: *Unprescaled minbias event rates for* $\sigma = 70\,\mathrm{mb}$ *and* $\mathcal{L} = 10^{31}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. *The quoted errors are only statistical.*

$10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Other tau trigger signatures as double tau 2tau35i have been studied by the ATLAS tau trigger community and they are foreseen to run in addition.

The fact that several physically interesting samples like Higgs, SUSY or $t\bar{t}$ are triggered by different tau trigger items is very convenient for different reasons:

- it leads to an increase of statistics;
- it is a robust approach against failures or inefficiencies of a particular trigger line (e.g. due to detector problems);
- it allows to reduce systematic uncertainties by comparing results of the same analysis repeated on samples selected with different trigger items.

Chapter 4

Offline Tau and $\mathrm{E}_{\mathrm{T}}^{miss}$ reconstruction

Many theoretical models, like Standard Model or SUSY at large $\tan\beta$*, predict Higgs bosons decaying more abundantly in tau leptons than in other lepton flavours. Despite this strong physics motivation to explore data with* τ *leptons in the final state, tau reconstruction at hadron colliders remains a very difficult task in terms of distinguishing interesting events from background processes dominated by QCD multi-jet production.*

A very good measurement of the missing transverse energy is also essential for many physics studies in ATLAS. Events with large $\mathrm{E}_{\mathrm{T}}^{miss}$ *are expected to be the key signature for new physics such as SUSY and extra dimensions. Moreover a good* $\mathrm{E}_{\mathrm{T}}^{miss}$ *measurement in terms of linearity and resolution is crucial for the efficient and accurate reconstruction of the Higgs boson mass when the Higgs boson decays into a pair of* τ *leptons.*

In this chapter two of the main algorithms developed in ATLAS for the tau reconstruction and for the $\mathrm{E}_{\mathrm{T}}^{miss}$ *measurement are described and their performance in terms of efficiency, rejection, linearity and resolution is shown.*

4.1 Hadronic τ Decays Reconstruction

Hadronically decaying τ leptons are distinguished from QCD jets on the basis of their shower shape in the calorimetric system, and of low track multiplicities contained in a narrow cone. Isolation from the rest of the event is required both in the calorimeters and the Inner Detector.

Calorimeters provide information on the energy deposit of the visible decay products (i.e. all decay products excluding neutrinos). Hadronically decaying τ leptons are well collimated, with an opening angle bounded by

the ratio m_τ/E_τ, leading to a relatively narrow shower in the EM calorimeter with a significant pure electromagnetic component for single-prong decays with one or few π^0s. On average in this case about 55% of the energy is carried by π^0s present among the decay products.

The Inner Detector provides information on the charged hadronic track or the collimated multi-track system. These tracks should neither match track segments in the muon spectrometer nor reveal features characteristic of an electron track (e.g. high threshold hits in the Transition Radiation Tracker). In the case of a multi-track system, the tracks are required to be well collimated in (η, ϕ) space and the invariant mass of the system should be below the τ lepton mass. The charge of the decaying τ lepton can be directly determined from the charge(s) of its decay product(s).

The calorimeter and tracking information should match, with narrow calorimeter cluster being found close to the track(s) impact point in the calorimeter.

At present, two complementary algorithms have been implemented into the ATLAS offline reconstruction software:

1. the calorimetry-based algorithm starts from clusters reconstructed in the hadronic and electromagnetic calorimeters and builds the identification variables based on information from the tracker and the calorimeter. The rejection power with respect to QCD jets comes almost entirely from the identification procedure and gives rejection rates of 100-10000 depending on the energy range, for an efficiency of about 30%;

2. the track-based algorithm starts from seeds built from few (low multiplicity) high quality tracks collimated around the leading one. The energy is calculated with an energy-flow algorithm based only on tracks and the energy in the electromagnetic calorimeter. All identification variables are built using information from the tracker and the calorimeters.

For the analysis discussed in this thesis, the calorimetry-based algorithm has been used to identify τs, because it has a better performance for high p_T τs. After a short overview of the features of τ lepton decays, a description of this offline reconstruction algorithm is given and performance results are shown.

4.1.1 τ Lepton in LHC Collisions

In the physics expected at the LHC the transverse momentum range of the τ leptons spans from below 10 GeV up to 500 GeV. The τ leptons decay hadronically in 64.8% of all cases, while in $\sim$ 17.8% (17.4%) of the cases they decay to an electron (muon) [58]. From the detection point of view, hadronic decay modes are divided according to the number of charged πs present among the decay products into single-prong (one charged π) and three-prong (three charged πs) decays. The small fraction (0.1%) of five-prong decays is usually too hard to detect in a hadronic collider environment. The $\tau \rightarrow \pi^{\pm}\nu$ mode contributes 22.4% to single-prong hadronic decays and the $\tau \rightarrow n\pi^0\pi^{\pm}\nu$ modes contributes 73.5%. For three-prong decays, the $\tau \rightarrow 3\pi^{\pm}\nu$ decay contributes 61.6%, and the $\tau \rightarrow n\pi^0 3\pi^{\pm}\nu$ mode 33.7%. In general, one- and three-prong modes are dominated by final states consisting of $\pi^{\pm}$ and π^0. There is a small percentage of decays containing $K^{\pm}$ which nevertheless can be identified using the same technique as for states with $\pi^{\pm}$ from the ATLAS detector point of view. A small percentage of states with K_S^0 cannot be easily classified as belonging to either the single-prong or three-prongs categories as the number of registered prongs depends on the actual K_S^0 interaction within the detector.

The lifetime of the τ lepton ($c\tau = 87.11\mu$m) in principle allows for the reconstruction of its decay vertex in the case of three-prong decays. The flight path in the detector increases with the Lorentz boost of the τ lepton, but at the same time the angular separation of the decay products decreases. A resulting transverse impact parameter of the τ decay products can be used to distinguish them from objects originating from the production vertex.

4.1.2 The Calorimeter-based Algorithm

For the reconstruction of hadronically decaying τ candidates, calorimeter clusters are used as seeds [59]. They are obtained from a sliding window clustering algorithm applied to so called calorimeter towers which are formed from cells of all calorimeter layers on a grid of size $\Delta\eta \times \Delta\phi = 0.1 \times 2\pi/64$. The energy and position are calculated from the clusters, while all cells with the full granularity of the corresponding calorimeters are used to calculate the quantities involved in τ identification as described in the following. Only clusters with a transverse energy $E_T > 15$ GeV are used. The probability for a true τ to be reconstructed as a cluster increases from 20% to 68% over the visible τ transverse energy range from 15 to 20 GeV and saturates at 98% for $E_T > 30$ GeV.

All cells within $\Delta R < 0.4$ around the barycenter of the cluster are then

calibrated with weights that are a function of the cell energy density and of the cell position[54]. These weights have been optimized for jets and only approximately for hadronic τ decays. The mean and width of a Gaussian fit to the ratio of the reconstructed and the generated τ decay product visible energy, $E_T^{\tau-vis}$, are shown in Figure 4.1 as a function of true $E_T^{\tau-vis}$ and η. The resolution is of the order of 10% and an offset in the range from +5 to -7% is observed in the τ energy range from 20 to 50 GeV, while at larger energies the offset is of the order of -3 to -5% [53].

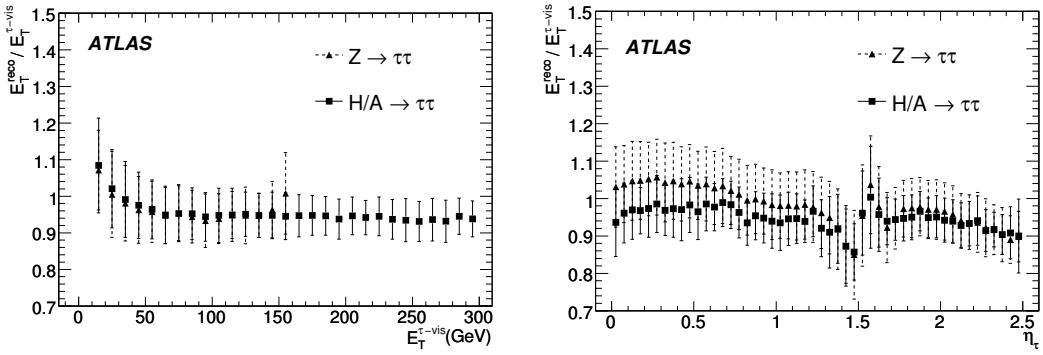

Figure 4.1: *The ratio of the reconstructed and the true* $E_T^{\tau-vis}$ *of the hadronic* τ *decay products is shown as a function of the true* $E_T^{\tau-vis}$ *(left) and* $|\eta|$ *(right) for* τ*s from* $Z \rightarrow \tau\tau$ *(triangles) and* $A \rightarrow \tau\tau$ *with* $m_A = 800$ *GeV (squares) decays.*

Several quantities that exploit the τ lepton properties have been combined in a likelihood function to discriminate hadronic τ decays from fake candidates originating from QCD jets. These quantities are described in detail in the following:

- **The electromagnetic radius** R_{em}:

  To exploit the smaller transverse shower profile in τ decays, the electromagnetic radius R_{em} is used, defined as

$$R_{\mathrm{em}} = \frac{\sum_{i=1}^{n} \mathrm{E}_{\mathrm{T},i} \sqrt{(\eta_i - \eta_{\mathrm{cluster}})^2 + (\phi_i - \phi_{\mathrm{cluster}})^2}}{\sum_{i=1}^{n} \mathrm{E}_{\mathrm{T},i}} \ , \qquad (4.1)$$

  where i runs over all cell in the electromagnetic calorimeter in a cluster with $\Delta R < 0.4$. The quantities η_i, ϕ_i, and $\mathrm{E}_{\mathrm{T},i}$ denote their position and transverse energy in cell i. Cells may have different sizes depending on the layer and their η value. The size varies from $\Delta\eta \times \Delta\phi = 0.003 \times 0.1$ in the η-strip region of the barrel to 0.025×0.025 for the second calorimeter layer. This leads to a dependence of the performance on η.

This variable shows good discrimination power at low E_T but becomes less effective at higher E_T.

- **Isolation in the calorimeter**:

  Clusters built from hadronic τ decays are well collimated, therefore rather tight isolation criteria can be used. Here a ring of $0.1 < \Delta R < 0.2$ was chosen as the isolation region and the quantity

  $$\Delta E_T^{12} = \frac{\sum_i E_{T,i}}{\sum_j E_{T,j}} , \tag{4.2}$$

  is calculated, where the indexes i and j run over all electromagnetic calorimeter cells in a cone around the cluster axis with $0.1 < \Delta R < 0.2$ and $\Delta R < 0.4$, respectively, and $E_{T,i}$ and $E_{T,j}$ denote the transverse cell energies.
  Like R_{em}, the ΔE_T^{12} distribution shows an E_T dependence and becomes narrower with increasing E_T. This variable also depends on the event type and is expected to be less effective for events with higher hadronic activity, like e.g. $t\bar{t}$ events.

- **Charge of the τ candidate**:

  The charge of a τ candidate is defined as the sum over the charge(s) of the associated track(s). The misidentification of the charge on the level of a few percent shows almost no E_T dependence.

- **Number of associated tracks**:

  The number of tracks, N_{tr}, associated with a given cluster within $\Delta R < 0.3$. The tracks are required to have $p_T > 2$ GeV, but no specific requirements on the quality of the track reconstruction is made. A significant fraction of events with zero, two, and even four tracks is observed for true hadronic τ decays.

- **Number of hits in the η strip layer**:

  The number of hits in η direction in the finely segmented strip detector, N_{strip}, in the first layer of the electromagnetic barrel calorimeter is also used in the likelihood discrimination. Cells in the η strip layer within $\Delta R < 0.4$ around the cluster axis are counted as hits if the energy deposited exceeds 200 MeV. In contrast to jets, a significant fraction of τ leptons deposit nearly no energy in the η strip layer ($\tau \to \pi\nu$ decays) and the number of corresponding hits is small.

- **Transverse energy width in the η strip layer**:

  The transverse energy width $\Delta\eta$ is defined as

$$\Delta\eta = \sqrt{\frac{\sum_{i=1}^{n} E_{T,i}^{strip} \left(\eta_i - \eta_{\text{cluster}}\right)^2}{\sum_{i=1}^{n} E_{T,i}^{strip}}} \ , \tag{4.3}$$

  where the sum runs over all strip cells in a cone with $\Delta R < 0.4$ around the cluster axis and $E_{T,i}^{strip}$ is the corresponding strip transverse energy. Like R_{em} it is a powerful discriminator at low E_T but it loses discrimination power with increasing E_T for collimated high E_T jets.

- **Lifetime signed pseudo impact parameter significance**:

  At present only a 2-dimensional impact parameter, also called the pseudo impact parameter d_0, is used. It is defined as the distance from the beam axis to the point of closest approach of the track in the plane perpendicular to the beam axis. From this information and from the τ-jet axis, a quantity denoted as lifetime signed pseudo impact parameter significance, defined as $\text{sig}_{d_0} = d_0/\sigma_{d_0}^2$ where σ is the impact parameter resolution, is calculated.

- E_T **over** p_T **of the leading track:** E_T/p_{T1} :

  A large fraction of the energy for τ decays is expected to be carried by the leading track and the ratio of the cluster energy E_T and the momentum of the leading track p_{T1} is expected to be close to 1. Values above one are expected from τ decay modes involving additional π^0s and for three-prong decays. This provides another discrimination against QCD jets, which are expected to have a more uniform distribution of p_T among the tracks. They are also expected to have more additional neutral particles. The E_T dependence is rather modest for τ decays but more pronounced for QCD jets, which tend to become more signal like with higher E_T.

In Figure 4.2 the distributions of a few discriminating variables are shown for signal and backgrounds for transverse cluster energies E_T in the range between 40 and 60 GeV and for candidates with 1 or 3 tracks.

For the calorimeter-based algorithm the τ identification is based on a one-dimensional likelihood ratio constructed from three discrete variables (N_{tr}, N_{strip} and the charge of the τ candidate) and five continuous variables (R_{em}, ΔE_T^{12}, $\Delta\eta$, sig_{d_0}, and E_T/p_{T1}). For the discrete variables the ratios are directly taken from the reference histograms. For the continuous variables fits

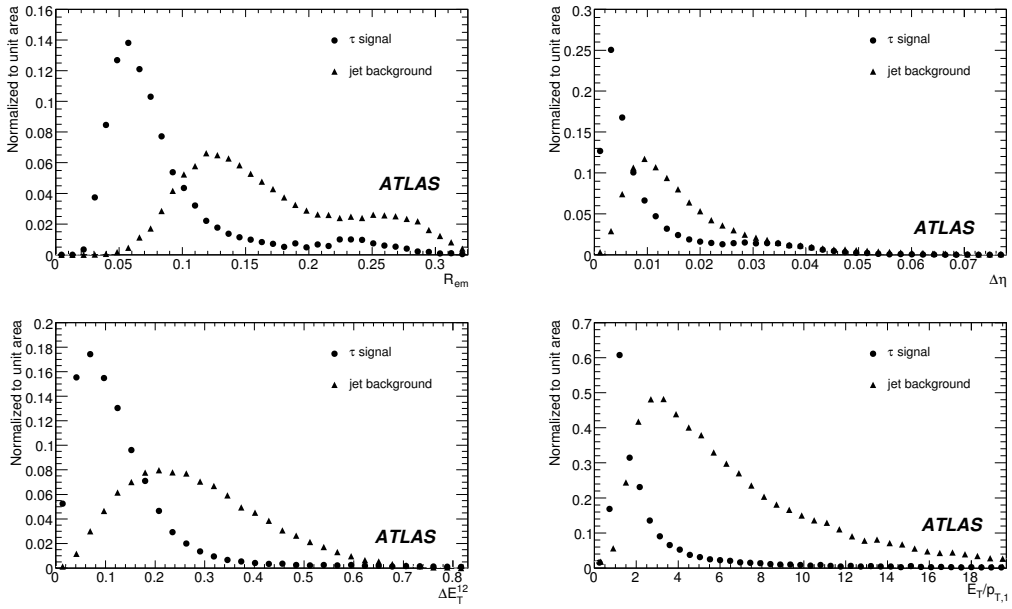

Figure 4.2: *The distributions of a few discriminating variables (electromagnetic radius and transverse energy width in the η strip layer on the top, energy isolation and* E_T/p_{T1} *of the leading track on the bottom) for true tau decays and jets with visible transverse cluster energies* E_T *in the range from 40 to 60 GeV and track multiplicities between 1 and 3.*

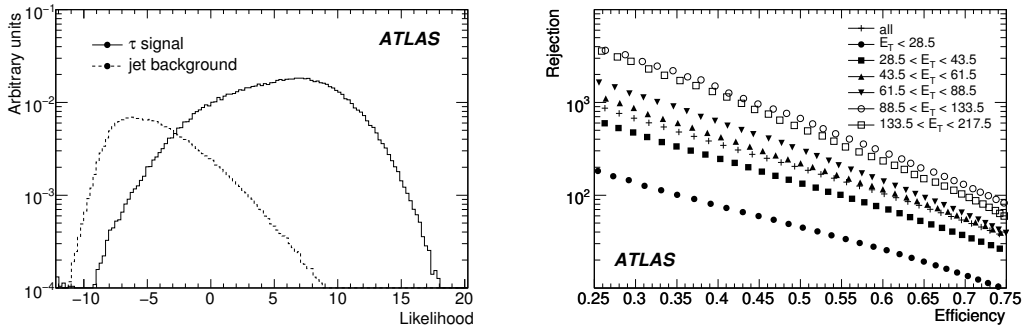

Figure 4.3: *The log likelihood (LLH) distribution for* τ *leptons (solid) and jets from QCD production (dashed). The likelihood is applied after a preselection on the number of associated tracks, i.e. requiring* $1 \leq N_{tr} \leq 3$ *(left). Efficiency for* τ *leptons and rejection against jets for different* E_T *ranges, achieved with the likelihood selection (right).*

of appropriate functions to each variable for all E_T bins have been performed. The distribution of the likelihood for taus and jets are shown in Figure 4.3. Despite any limitation from using only one-dimensional distributions it shows a good separation power.

It should be noted that the choice of the τ identification efficiency depends on the physics channel.

The expected performance of the τ reconstruction algorithm is illustrated in Figure 4.4, where the the jet rejection versus the τ efficiency is shown for one and three-prong hadronic τ-decays and for two different ranges of the visible transverse energy. The jet rejections are computed with respect to jets reconstructed from true particles in the Monte Carlo. The rejections obtained are a factor of ten higher for one-prong decays than for three-prong decays. For an efficiency of 30% for one-prong decays, the rejection against jets is typically between 1000 and 5000, as illustrated more quantitatively in Table 4.1 as a function of the visible transverse energy.

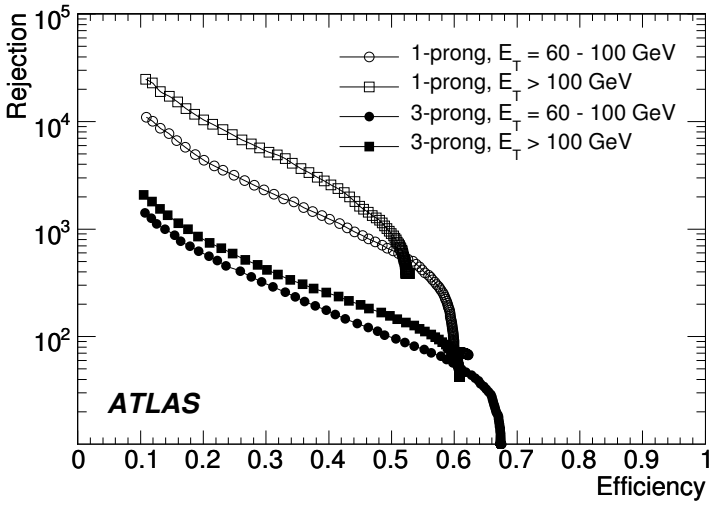

Figure 4.4: *Expected performance for the calorimeter-based algorithm with the likelihood selection. The rejection rates against jets from Monte-Carlo particles as a function of the efficiency for hadronic τ decays for various ranges of the visible transverse energy are shown. For signal events $Z \to \tau\tau$ and $bbH, H \to \tau\tau$ with $m_H = 800$ GeV were used, for the background QCD dijet samples were used.*

Algorithm	E_T = 30-60 GeV	E_T = 60-100 GeV	E_T > 100 GeV
Calo-based (likelihood)	1p: 1130 ± 50 3p: 187 ± 3	1p: 2240 ± 140 3p: 310 ± 7	1p: 4370 ± 280 3p: 423 ± 8

Table 4.1: *Rejection against jets for a 30% efficiency, for the one-prong (1p) and three-prong (3p) candidates. The efficiencies are normalized to true hadronic τ decays. For the signal $Z \rightarrow \tau\tau$ events and events from bbH, $H \rightarrow \tau\tau$ with $m_H = 800$ GeV were used; for the background QCD dijet-samples were used. The errors given are statistical only.*

4.2 E_T^{miss} Measurement

An important requirement on the measurement of E_T^{miss} is to minimize the impact of limited detector coverage, finite detector resolution, presence of dead regions and different sources of noise that produce fake E_T^{miss}. The ATLAS calorimeter coverage extends to large pseudo-rapidity to minimize the impact of high energy particles escaping in the very forward direction. Nevertheless, there are inactive transition regions between different calorimeters that produce fake E_T^{miss}. Dead and noisy readout channels in the running detector, if present, will also produce fake E_T^{miss}. Such sources of fake E_T^{miss} can significantly enhance the background from $Z \rightarrow jj$ events accompanied by high-p_T jets in Higgs boson searches when the Higgs boson decays into two τ-jets and neutrinos as it will be seen in the following chapter.

The transverse missing energy in ATLAS is primarily reconstructed from energy deposits in the calorimeter and reconstructed muon tracks [57]. The calorimeters play a crucial role in the E_T^{miss} measurement and an important first step is the suppression of their noise. Apart from the hard scattering process of interest, many other sources, such as the underlying event, multiple interactions, pileup and coherent electronic noise, lead to energy deposits and/or muon tracks. Classifying the energy deposits into various types (e.g. electrons or jets) and calibrating them accordingly is the essential key for an optimal E_T^{miss} measurement. There are two algorithms for E_T^{miss} reconstruction in ATLAS that emphasize different aspects of energy classification and calibration.

- The cell-based algorithm starts from the energy deposits in calorimeter cells that survive a noise suppression procedure. The cells can be calibrated using global calibration weights depending on their energy density. This procedure will be robust already at initial data taking because it does not rely on other reconstructed objects. In a subsequent step, the cells can be calibrated according to the reconstructed object

they are assigned to. Corrections are applied for the muon energy and for the energy lost in the cryostat.

- The object-based algorithm starts from the reconstructed, calibrated and classified objects in the event. The energy outside these objects is further classified as low p_T deposit from charged and neutral pions and calibrated accordingly.

The noise suppression in the calorimeter is common for the cell- and object-based algorithm and is described below, followed by a detailed description of the cell-based algorithm that is the algorithm used to perform the analysis described in this thesis work.

4.2.1 Calorimeter Noise Suppression

The electronic noise alone in the ∼200k readout channels of the ATLAS calorimeters contributes about 13 GeV to the width of the E_T^{miss} distribution. Especially in events that do not have large E_T^{miss}, the noise suppression is of crucial importance. For the E_T^{miss} measurement, two noise suppression methods have been studied so far. Both require knowledge of the width of the noise distribution, σ_{noise}, which can be either purely electronics noise or a combination of electronics and pileup noise.

Standard Noise Suppression Method. This method only uses calorimeter cells with energies larger than a threshold, generally corresponding to a certain multiple of σ_{noise}. The threshold is optimized for E_T^{miss} resolution, for the total transverse energy in the calorimeters, ΣE_T, and for the highest p_T jet to be close to the case without noise simulation. Two cases are studied: a symmetric threshold ($|E_{cell}| > n \times \sigma_{noise}$) and an asymmetric one ($E_{cell} > n \times \sigma_{noise}$). A symmetric threshold with n = 2 for all calorimeters is generally used.

Noise Suppression using TopoClusters. This second method only uses cells in 3-dimensional topological calorimeter clusters [60] [61], hereafter called TopoClusters. A TopoCluster is reconstructed starting with a seed cell with an absolute energy value $|E_{cell}| > 4\sigma_{noise}$ to which all neighbouring cells are added. For those added cells that have a $|E_{cell}| > 2\sigma_{noise}$, the neighbouring cells are also added to the cluster. Finally the cells at the boundary are required to have $|E_{cell}| > 0\sigma_{noise}$. The cells that constitute the TopoCluster are hereafter called TopoCells. This set of thresholds, referred to as 4/2/0, is optimized to suppress electronics noise as well as pileup from minimum bias events, while keeping the single pion efficiency as high as possible. As a result of the large energy density of electromagnetic showers, the

π^0 reconstruction efficiency is high (close to 100% for energy > 4 GeV) for the 4/2/0 configuration. On the other hand, the reconstruction efficiency for charged pions is very sensitive to the parameters of the TopoClusters. For example, changing the cuts on neighbors from 4/2 to 6/3, the $\pi^{\pm}$ efficiency significantly decreases for TopoClusters with E < 4 GeV. This sensitivity highlights the importance of a good modeling of the noise from first data.

In $Z \rightarrow \nu\nu$ events simulated with electronics noise, the E_T^{miss} resolution degrades by only 3% for the 4/2/0 configuration as compared to the same events without noise added. Also, the TopoCluster algorithm performs better in terms of linearity and resolution of the E_T^{miss} measurement, compared to the standard noise suppression method. For this reasons the noise suppression method chosen in the cell-based E_T^{miss} algorithm is based on TopoClusters 4/2/0.

4.2.2 E_T^{miss} Cell-based Reconstruction Algorithm

The E_T^{miss} reconstructed with the cell-based algorithm includes contributions from transverse energy deposits in the calorimeters, corrections for energy loss in the cryostat and measured muons [57]:

$$E_{x,y}^{miss,Final} = E_{x,y}^{miss,Calo} + E_{x,y}^{miss,Cryo} + E_{x,y}^{miss,Muon} \ . \tag{4.4}$$

A detailed description of the three terms of the previous equation, referred to as calorimeter, cryostat and muon terms, is given in some detail in the following.

E_T^{miss} Calorimeter Term ($E_{x,y}^{miss,Calo}$)

As described in the previous section, the first step is to select calorimeter cells that belong to reconstructed TopoClusters to minimize the impact of noise. The x and y components of the calorimeter E_T^{miss} term are calculated from the transverse energies measured in TopoCells:

$$E_{x,y}^{miss,Calo} = -\sum_{TopoCells} E_{x,y} \ . \tag{4.5}$$

The total transverse energy in the calorimeters, ΣE_T, is calculated from the scalar sum of E_T of all TopoCells:

$$\Sigma E_T^{Calo} = \sum_{TopoCells} E_T \ . \tag{4.6}$$

The straightforward result, obtained by using the electromagnetic calibration for all cells, gives a large shift in the E_T^{miss} scale with respect to true

E_T^{miss} . This result illustrates the necessity of developing a dedicated calibration scheme to reduce this systematic E_T^{miss} scale shift and optimize its resolution. This goal is achieved in several steps according to the origin of the cell signal. The calibration term depends on whether the energy deposits in the calorimeter are electromagnetic or hadronic in nature and whether they are associated with high p_T particles.

To classify energy deposits, schemes to calibrate hadronic showers such as 'Global' calibration or 'Local-Hadronic' calibration [54] utilize the energy density in a cell. Electromagnetic showers tend to have higher energy densities as compared to hadronic showers. The Local-Hadronic calibration scheme uses further information related to shape and depth of the calorimetric shower to classify a TopoCluster. The next step in E_T^{miss} reconstruction is to globally calibrate all calorimeter cells using the Global or Local-Hadronic calibration schemes.

E_T^{miss} Cryostat Term ($E_{x,y}^{miss,Cryo}$)

Hadronic showers can lose energy in the cryostat which is located between the LAr barrel electromagnetic calorimeter and the Tile barrel hadronic calorimeter. The cryostat is about half an interaction length thick. The E_T^{miss} reconstruction recovers this loss of energy in the cryostat using the correlation of energies between the last layer of the LAr calorimeter and the first layer of the hadronic calorimeter. A similar correction is also applied for the end-cap cryostats. This correction is called the cryostat term when used for jet energy correction [54]. It is defined as follows:

$$E_{x,y}^{miss,Cryo} = -\sum_{recJets} E_{jet_{x,y}}^{Cryo} , \tag{4.7}$$

where all reconstructed jets are summed in the event, and

$$E_{jet}^{Cryo} = w^{Cryo}\sqrt{E_{EM3} \times E_{HAD}} , \tag{4.8}$$

where w^{Cryo} is a calibration weight determined together with the cell calibration weights in the Global calibration, and E_{EM3} and E_{HAD} are the jet energies in the third layer of the electromagnetic calorimeter and in the first layer of the hadronic calorimeter, respectively. The cryostat correction turns out to be non-negligible for $high-p_T$ jets. It contributes at the level of $\sim$5% for jets with p_T above 500 GeV.

E_T^{miss} Muon Term ($E_{x,y}^{miss,Muon}$)

The E_T^{miss} muon term is calculated from the momenta of muons measured in a large range of pseudorapidity, defined by $|\eta| < 2.7$:

$$E_{x,y}^{miss,Muon} = - \sum_{RecMuons} E_{x,y} \ . \tag{4.9}$$

In the region $|\eta| < 2.5$ only good-quality muons in the muon spectrometer with a matched track in the Inner Detector are considered. The matching requirement reduces considerably the contributions from fake muons, sometimes created from high hit multiplicities in the muon spectrometer in events with very energetic jets. For higher values of the pseudo-rapidity ($2.5 < |\eta| < 2.7$), outside the fiducial volume of the Inner Detector, there is no matched track required and the muon spectrometer alone is used. The muon energy measured by the muon spectrometer is taken in the two cases. The energy lost in the calorimeter is already included in the calorimeter term. No p_T threshold cut is applied to reconstructed muons. Apart from the loss of muons outside the acceptance of the muon spectrometer ($|\eta| < 2.7$), there is a loss of muons in other regions of the detector due to limited coverage of the muon spectrometer. The muons reconstructed from the Inner Detector and the calorimeter energy deposits could in principle be used to recover these events, but here they are not used for this purpose yet.

Refined Calibration of E_T^{miss} Calorimeter Term

The final step is the refinement of the calibration of cells associated with each $high-p_T$ object. Calorimeter cells are associated with a parent reconstructed and identified $high-p_T$ object, in a chosen order: electrons, photons, muons, hadronically decaying τ-leptons, b-jets and light jets. Refined calibration of the object is then used to replace the initial global calibration of the cells. The calibration of these objects is known to higher accuracy than the global calibration, enabling to improve the E_T^{miss} reconstruction. The calorimeter cells are associated with the reconstructed objects through the use of an association map. This map is filled starting from the reconstructed/identified objects in the chosen order, navigating back to their component clusters and back again to their cells. If a cell belongs to several kinds of reconstructed objects, only the first association is included in the map, i.e. the overlap removal is done at cell level. This avoids double counting of cells in the E_T^{miss} calculation. If a cell belongs to more than one object of the same kind, all associations are included in the map and the geometrical weight of the cells, accounting for the sharing of energy of cells owned by to two different TopoClusters, are also included to avoid double counting.

Attention has to be paid to the calibration of cells inside different objects. For example, for electrons/photons, the final cluster-level calibration (which can be propagated back to the cell-level) corrects for upstream material, longitudinal leakage and out-of-cone energy. The last correction should not be applied in the E_T^{miss} calculation because the contribution of cells outside objects already accounts for it. In a similar way, for τ lepton decays and for jets, the overall scale factors which correct the energy for physics effects like final state radiation, fragmentation or the underlying event as well as for the effects due to the clustering algorithm are not applied in the calculation of E_T^{miss} , because they also contain the out-of-cluster correction. All TopoCells, even if not associated with any $high-p_T$ reconstructed object, are used in the E_T^{miss} calculation. They are calibrated using the global calibration scheme. The importance of the energy deposits of these cells for the E_T^{miss} calculation is shown in Figure 4.5 . The shift in the absolute value of the reconstructed E_T^{miss} increases by about 1 GeV while the resolution is degraded by a factor $\sim$1.25. Once the cells are associated with categories of objects as described

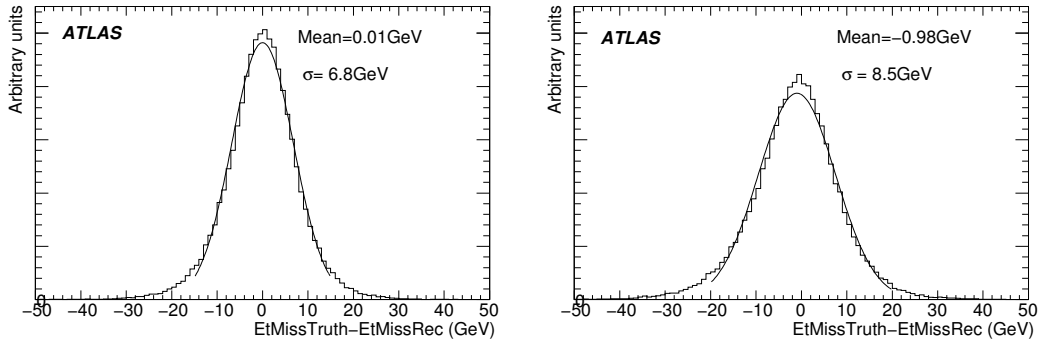

Figure 4.5: *Distribution of the difference between true and reconstructed* E_T^{miss} *for* $Z \rightarrow \tau^+\tau^-$ *events with (left) and without (right) the inclusion of cells in TopoClusters not associated with reconstructed high* $-$ p_T *objects.*

above, the contribution to E_T^{miss} is calculated as:

$$E_{x,y}^{miss,Calo} = E_{x,y}^{miss,RefCalib} = -(E_{x,y}^{miss,RefEle} + E_{x,y}^{miss,RefTau} + \\ +E_{x,y}^{miss,Refbjets} + E_{x,y}^{miss,RefJets} + E_{x,y}^{miss,RefMuo} + E_{x,y}^{miss,RefOut}) \ , \quad (4.10)$$

where each term is calculated from the negative of the sum of calibrated cells inside a specific object and the $E_{x,y}^{miss,RefOut}$ is calculated from the cells in TopoClusters which are not included in the reconstructed objects. In the following the final E_T^{miss} calculation obtained from Equation 4.4 with $E_{x,y}^{miss,Calo} = E_{x,y}^{miss,RefCalib}$ will be used.

4.2.3 Performance of Reconstructed E_T^{miss}

The performance of the E_T^{miss} reconstruction algorithm can be assessed in terms of linearity and resolution of the reconstructed E_T^{miss} as a function of the true missing transverse energy, $E_T^{miss,True}$.

The E_T^{miss} linearity is defined by the following expression:

$$Linearity = (|E_T^{miss,True}| - |E_T^{miss}|)/|E_T^{miss,True}| \ , \tag{4.11}$$

where E_T^{miss} and $E_T^{miss,True}$ are reconstructed and true E_T^{miss} , respectively. Figure 4.6 [57] shows the linearity as a function of $E_T^{miss,True}$ for a number of different physics processes. The following statements summarize the behavior of the linearity distributions:

- the uncalibrated E_T^{miss} corresponds to the use of cell energies at the electromagnetic scale and shows a large systematic bias of 30%. In $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ decays, the bias is smaller since the hadronic activity on average is smaller;

- the reconstructed E_T^{miss} based on globally calibrated cell energies and reconstructed muons shows deviations from linearity at the level of 5%;

- the reconstructed E_T^{miss} including the cryostat correction is linear within 1% for all processes except for $W \rightarrow e\nu$;

- the refined E_T^{miss} calibration, which optimizes the calibration using the reconstructed object identity, recovers the linearity for $W \rightarrow e\nu$ events to within 1%. The refined calibration also gives the best resolution when compared with the above steps of calibration.

The linearity for $A \rightarrow \tau^+\tau^-$ with $m_A = 800$ GeV is shown in Figure 4.7 [57] as a function of $E_T^{miss,True}$. The bias of linearity at low $E_T^{miss,True}$ is due to the finite resolution of the E_T^{miss} measurement. The reconstructed E_T^{miss} is positive by definition, so the linearity is negative when the true E_T^{miss} is near to zero. Excluding the events with $E_T^{miss,True} < 40$ GeV, which have a small statistics, the observed linearity is found to be within 2%.

The resolution is estimated from the width of the $E_{x,y}^{miss} - E_{x,y}^{miss,True}$ distribution in bins of the total transverse energy deposited in the calorimeters (ΣE_T). The core of each distribution is fitted to a Gaussian to estimate its width. Figure 4.8 [57] shows the σ of the fit plotted as a function of ΣE_T when the refined calibration is applied. The E_T^{miss} resolution approximately follows a stochastic behaviour as a function of ΣE_T. Deviations from this simple behaviour are expected, and observed for low values of ΣE_T where

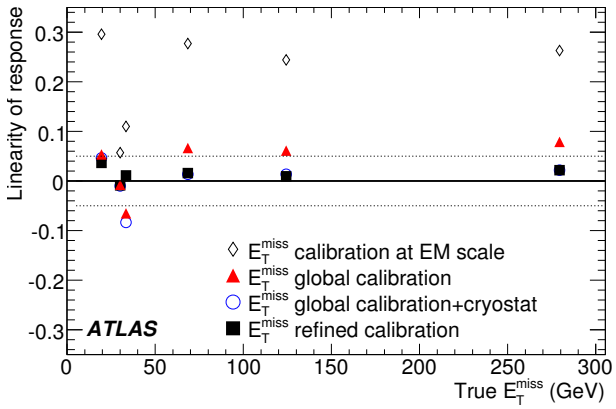

Figure 4.6: *Linearity of response for reconstructed* E_T^{miss} *as a function of the average true* E_T^{miss} *for different physics processes and for the different steps of* E_T^{miss} *reconstruction. The points at average true* E_T^{miss} *of 20 GeV are from* $Z \rightarrow \tau^+\tau^-$ *events, those at 35 GeV are from* $W \rightarrow e\nu$ *and* $W \rightarrow \mu\nu$ *events, those at 68 GeV are from semi-leptonic* $t\bar{t}$ *events, those at 124 GeV are from* $A \rightarrow \tau^+\tau^-$ *events with* $m_A = 800$ *GeV, and those at 280 GeV are from events containing supersymmetric particles at a mass scale of 1 TeV.*

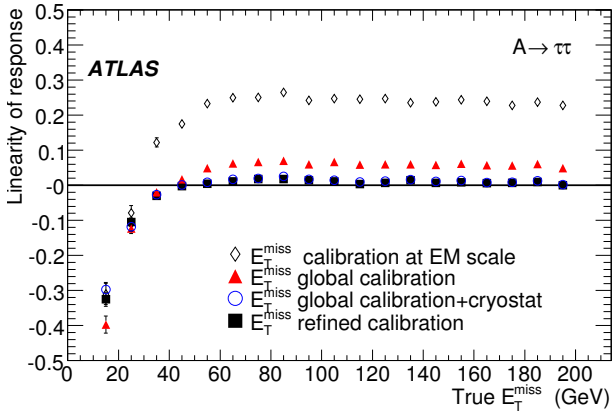

Figure 4.7: *Linearity of response for reconstructed* E_T^{miss} *as a function of the true* E_T^{miss} *for* $A \rightarrow \tau^+\tau^-$ *events with* $m_A = 800$ *GeV.*

the contribution of noise is important and for very high values of ΣE_T where the constant term in the resolution of the calorimetric energy measurement dominates.

The E_T^{miss} resolution is fitted with a function $\sigma = a \cdot \sqrt{\Sigma E_T}$ for values of ΣE_T between 20 and 2000 GeV. The parameter a, which quantifies the E_T^{miss} resolution, varies between 0.53 and 0.57 (see Figure 4.8 left and right, respectively).

A good performance in terms of linearity and resolution may enhance the

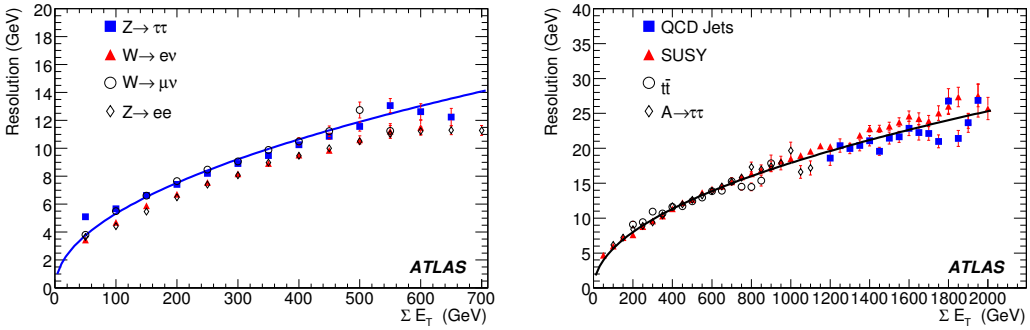

Figure 4.8: *Resolution of the two E_T^{miss} components with refined calibration as a function of the total transverse energy ΣE_T, for low to medium values (left) and for higher values (right). The curves correspond to the best fits $\sigma = 0.53 \cdot \sqrt{\Sigma E_T}$, through the points from $Z \rightarrow \tau^+\tau^-$ events (left) and $\sigma = 0.57 \cdot \sqrt{\Sigma E_T}$, through the points from $A \rightarrow \tau^+\tau^-$ events (right). The points from $A \rightarrow \tau^+\tau^-$ events are for masses m_A ranging from 150 to 800 GeV and the points from QCD jets correspond to di-jet events with 560< p_T < 1120 GeV.*

capability to reconstruct the mass of final states which involve neutrinos. Despite the presence of several neutrinos in the final state, the invariant mass of the $\tau\tau$ pair can be reconstructed in $Z \rightarrow \tau^+\tau^-$ and supersymmetric Higgs boson decays $A \rightarrow \tau^+\tau^-$ under simplifying assumptions [62][63]. Figure 4.9 [57] shows the reconstructed mass peaks of $Z \rightarrow \tau^+\tau^-$ and supersymmetric Higgs boson $A \rightarrow \tau^+\tau^-$ decays with $m_A = 450$ GeV. The reconstructed masses are correct to $\sim 2\%$ and the mass resolution is approximately 11%. Nevertheless, significant tails remain in the distributions because of the highly non-Gaussian effects induced by mis-measurements of E_T^{miss} and by the approximations used.

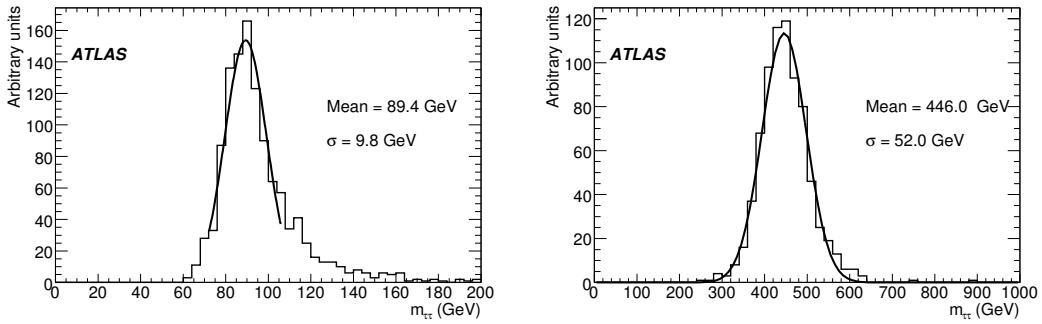

Figure 4.9: *Distributions of the reconstructed invariant mass of τ-lepton pairs with one τ-lepton decaying to a lepton and the other one decaying to hadrons. The results are shown for $Z \rightarrow \tau^+\tau^-$ decays (left) and for $A \rightarrow \tau^+\tau^-$ decays with m_A = 450 GeV (right).*

Chapter 5

The SM Higgs boson decay into τ-leptons

In ATLAS, reasonable trigger rates together with high efficiencies have been obtained using the tau trigger in association with $\mathrm{E}_{\mathrm{T}}^{miss}$*. At the offline level, the ATLAS algorithms for the tau reconstruction provide a very good rejection against QCD jets and an accurate measurement of the* $\mathrm{E}_{\mathrm{T}}^{miss}$ *allows to select events with neutrinos in the final state and to reconstruct the invariant mass of particles decaying into a* τ*-pair with a very good linearity and resolution. This allows to search for the Standard Model Higgs decaying into* τ*-pairs, not only in the fully leptonic (ll) and semi-leptonic (lh) decay modes of the* τ*-pair, but also in the fully hadronic (hh) final state. The Vector Boson Fusion (VBF) production mechanism for the Higgs boson is exploited for its distinctive topological characteristics.*

5.1 Signal and Background

As already mentioned in the section 1.4, at the LHC the production cross-section for a Standard Model Higgs boson is dominated by the gluon-gluon fusion (Figure 1.6, page 20). The second largest production mechanism is the fusion of vector bosons radiated off from initial-state quarks. Although in the low mass region the VBF production of the Higgs is reduced by a factor 5, it offers some peculiar kinematic features which can be exploited to extract the signal from the background. As the Feynman diagram in Figure 5.1 shows, the signal is characterized by the presence of two jets, produced by the fragmentation and hadronization of the associated quarks. They are expected to be very energetic and emitted close to the beam axis, with a large η separation. Moreover, because of the lack of color exchange

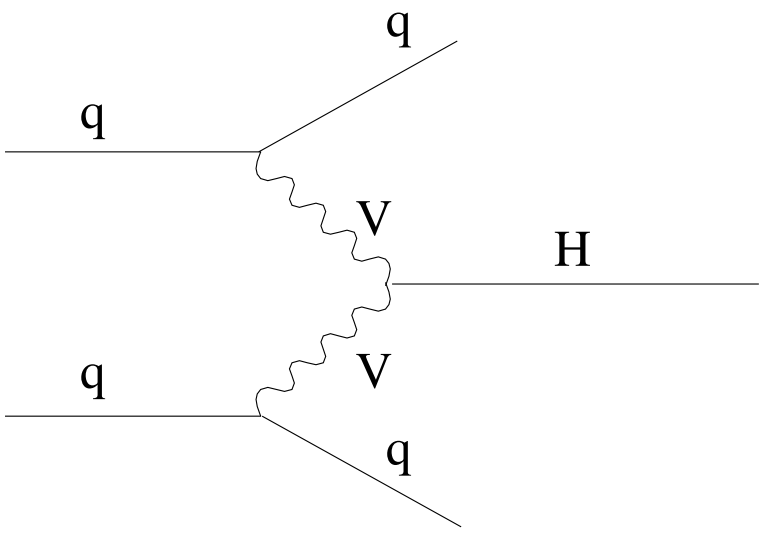

Figure 5.1: *Feynman diagram for the Vector Boson Fusion production.*

between quarks in the initial state, a low hadronic activity is expected in the central region of the detector. These peculiar kinematic signatures of VBF Higgs production allow for a signal to background separation good enough to extract the $H \rightarrow \tau\tau$ signal. In the Table 5.1 the cross sections for different Higgs masses are shown for the VBF production mechanism, together with the $H \rightarrow \tau\tau$ branching ratio. The fully hadronic final state has a branching ratio of 42%.

Higgs Mass (GeV)	105	110	115	120	125	130	135
$\sigma(qqH)$ (pb)	5.0	4.81	4.62	4.47	4.26	4.13	3.99
$\sigma(qqH) \cdot BR(H \rightarrow \tau\tau)$ (fb)	394.7	372.0	341.8	309.1	266.8	225.4	180.1

Table 5.1: *NLO cross sections for Higgs produced via Vector Boson Fusion and decaying into $\tau^+\tau^-$ for different Higgs boson masses.*

The most important irreducible backgrounds to this analysis comes from the QCD $Z + jets$ production and electroweak (EW) $Z + jets$ production via vector boson fusion, with Z decaying into a tau-pair. The QCD $Z + jets$ cross section is about a factor $6{\times}10^3$ larger than that one of the signal. This requires very effective kinematic cuts and a narrow Z reconstructed invariant mass to extract the signal, in particular for low Higgs boson masses.

The EW $Z + jets$ production has a factor six higher cross section, but in this case the kinematic cuts are not expected to be very effective.

The major source of reducible background for the fully hadronic final state of $H \rightarrow \tau\tau$ comes from the QCD multijets. Having a cross section of

several mb, a factor of at least 10^{11}–10^{12} must be reached in the rejection to extract the signal.

Also the $t\bar{t}$ production, and $W + jets$ production, have to be considered as backgrounds. The first one has an inclusive cross section $\sigma_{t\bar{t}} = 833$ pb and its events can contain two, one or no τs depending on the decay of the two Ws, produced in $t \to Wb$ decay. The second background has an inclusive production cross section σ_{W+jets}= 20510 pb. In this case one τ can be produced in the $W \to \tau\nu$ decay and an additional τ can be faked by an ordinary jet.

5.2 Monte Carlo Samples

To estimate the sensitivity of ATLAS the most recent Monte Carlo tools are needed. The most challenging aspect of the theoretical calculations is the description of jet activity. The signal samples were produced with HERWIG [64] and PYTHIA [29]. The QCD Z+jets and W+jets samples were produced with ALPGEN [30], which employs the MLM matching [65] between the hard process calculated with a leading-order matrix element for up to 5 jets and the parton shower of HERWIG. The electroweak Z+jets background was simulated with SHERPA [66]. The $t\bar{t}$+jets background sample was generated with MC@NLO [67]. The relevant differences among the generators are explained in [68]. In all processes with taus, the tau decay was simulated using TAUOLA [69]. Additional photon radiation from charged leptons was simulated with PHOTOS [70]. The production cross-section for the signal is based on the next-to-leading order (NLO) calculation and the k-factor (the ratio of the cross-section to that predicted by the lowest order calculation) is around 5% in the tau pair mass range of 100-150 GeV [71].

Because the GEANT-based detector simulation is computationally intensive, an event filter was applied to each sample after the parton shower and hadronization. Most processes were required to have at least one lepton in the final state. The filter bias has been studied and well-validated, but it affects our ability to estimate background rates early in the analysis flow. For this reason, a significant Monte Carlo sample was produced with the ATLAS fast simulation, ATLFAST, without any event filter [72]. These ATLFAST samples are used for systematic studies and to aid in the estimation of background rates.

The effect of in-time pileup (i.e. other soft pp collisions in the same bunch crossing), out-of-time pileup (i.e. pp collisions in neighboring bunch crossings), and the underlying event (i.e. multi-parton scattering and soft activity in the pp collision of interest) are all important for this analysis.

Process	Event Filter (ε)	σ(pb)	Events
HERWIG VBF $H(105) \to \tau\tau \to hh$	HIGGS_HH (0.230)	0.091	28000
HERWIG VBF $H(110) \to \tau\tau \to hh$	HIGGS_HH (0.234)	0.087	33000
HERWIG VBF $H(115) \to \tau\tau \to hh$	HIGGS_HH (0.238)	0.081	29250
HERWIG VBF $H(120) \to \tau\tau \to hh$	HIGGS_HH (0.237)	0.073	92000
HERWIG VBF $H(125) \to \tau\tau \to hh$	HIGGS_HH (0.241)	0.064	29750
HERWIG VBF $H(130) \to \tau\tau \to hh$	HIGGS_HH (0.246)	0.055	32500
HERWIG VBF $H(135) \to \tau\tau \to hh$	HIGGS_HH (0.247)	0.044	29000
Pileup L=10^{33}cm^{-2}s^{-1} VBF $H(120) \to \tau\tau \to hh$	HIGGS_HH (0.237)	0.073	38500
SHERPA ELWK $Z \to \tau\tau$+2,3p		1.693	185750
ALPGEN $Z \to \tau\tau \to hh$+1p	VBF_HH2 (0.0765)	17.7	83250
ALPGEN $Z \to \tau\tau \to hh$+2p	VBF_HH2 (0.196)	15.2	94750
ALPGEN $Z \to \tau\tau \to hh$+3p	VBF_HH2 (0.219)	5.30	142750
ALPGEN $Z \to \tau\tau \to hh$+4p	VBF_HH2 (0.224)	1.54	89000
ALPGEN $Z \to \tau\tau \to hh$+5p	VBF_HH2 (0.238)	0.60	72000
MC@ NLO $t\bar{t}$ (no hadronic)	ONELEP (0.599)	498.9	1012940
ALPGEN $W \to \tau\nu$+1p	VBF_HH1 (0.0943)	284	48250
ALPGEN $W \to \tau\nu$+2p	VBF_HH1 (0.449)	428	89500
ALPGEN $W \to \tau\nu$+3p	VBF_HH1 (0.543)	149.5	77250
ALPGEN $W \to \tau\nu$+4p	VBF_HH1 (0.559)	43.5	24000
ALPGEN $W \to \tau\nu$+5p	VBF_HH1 (0.557)	16.0	6000
QCD dijets (full)		$\sim 1.4 \times 10^9$	1,503,250
QCD dijets (fast)		$\sim 1.4 \times 10^9$	80,000,000

Table 5.2: *Used MC samples for signal and backgrounds. The reported cross sections σ(pb) already take into account the efficiency of the event filter applied.*

Event Filter	description
HIGGS_HH :	N_{τ_h}=2, $p_{T\tau_h} \geq 12$ GeV, $\lvert\eta_{\tau_h}\rvert \leq 2.7$, $\Delta\phi_{\tau_h\tau_h} \leq 2.9$
	$N_{jet} \geq 1$, $p_{Tjet} \geq 20$ GeV,
	$\lvert\eta_{jet}\rvert \leq 5.0$ for truth jet(Cone4TruthJets),
	$p_{Tjet} \geq 40$ GeV for highest p_T jet
VBF_HH1 :	N_{τ_h}=1, $p_{T\tau_h} \geq 12$ GeV, $\lvert\eta_{\tau_h}\rvert \leq 2.7$,
	$N_{jet} \geq 2$, $p_{Tjet} \geq 20$ GeV,
	$\lvert\eta_{jet}\rvert \leq 5.0$ for truth jet(Cone4TruthJets),
	$p_{Tjet} \geq 40$ GeV for highest p_T jet
VBF_HH2 :	N_{τ_h}=2, $p_{T\tau_h} \geq 12$ GeV, $\lvert\eta_{\tau_h}\rvert \leq 2.7$, $\Delta\phi_{\tau_h\tau_h} \leq 2.9$
	$N_{jet} \geq 2$, $p_{Tjet} \geq 15$ GeV,
	$\lvert\eta_{jet}\rvert \leq 5.0$ for truth jet(Cone4TruthJets),
	$p_{Tjet} \geq 20$ GeV for highest p_T jet
ONELEP :	$N_{e \text{ or } \mu \text{ or } \tau} \geq 1$, $p_{T\, e \text{ or } \mu \text{ or } \tau} \geq 10$ GeV, $\lvert\eta_{e \text{ or } \mu \text{ or } \tau}\rvert \leq 2.7$

Table 5.3: *Summary of Event Filters used to produce signal and background samples.*

The underlying event has a substantial theoretical uncertainty, and different model predictions for the underlying event activity vary by large factors when extrapolating to the LHC energy range. Fortunately, the underlying event activity will be one of the first measurements at the LHC. The pileup interactions are incorporated early in the simulation chain, at the time when the detector readout is simulated.

In Table 5.2 the list of the samples that have been used in this analysis is shown. A description of the event filters can be found in Table 5.3 [73] [74] [75]. The η cuts that have been applied to the τs and jets are justified by the fact that the first ones can be reconstructed within the region covered by the Inner Detector ($|\eta| < 2.5$) and the second ones within the acceptance of the calorimetric system ($|\eta| < 4.9$). The ϕ and p_T requirements are related to the analysis procedure that will be described in the following. For many background samples, the available statistics corresponds to few fb^{-1} of integrated luminosity. To estimate their contribution to the final number of events after the analysis selection, which is foreseen to be done with an integrated luminosity of 30 fb^{-1}, a cut-factorization procedure has been applied. For the multijet QCD background, a sufficient MC statistics cannot be reached even in fast simulation. To estimate this important background a weighting procedure based on the τ rejection parametrization obtained from fully simulated data has been applied.

Trigger menu	Efficiency × Acceptance(%)
`tau35i_xe40`	3.67±0.02

Table 5.4: *The product of efficiency and acceptance for the signal from the* `tau35i_xe40` *trigger.*

5.3 Events selection

5.3.1 Triggering

The ATLAS trigger system provides several possibilities which take advantage of the signal's complex final state. The entire trigger chain has been simulated with the use of the current trigger algorithms and trigger menus. We use `tau35i_xe40` as the primary trigger menu for the hh-channel. The expected trigger acceptance of `tau35i_xe40` is listed in Table 5.4. The trigger efficiency for the signal events (for $m_H = 120$ GeV) is 3.7% for `tau35i_xe40`. The disadvantage of the missing E_T trigger is the relatively low efficiency on signal; therefore, alternative menus like double tau menus are now being developed.

5.3.2 Hadronic tau identification

As already explained in the section 4.1, ATLAS employs two algorithms for hadronic tau reconstruction. The calorimeter-seeded algorithm was used for this analysis. The calorimeter-seeded algorithm provides a log-likelihood ratio (LLH) that exploits track quality and shower shape information to discriminate between taus and jets [53]. The discriminating variable is designed to maintain a high tau efficiency while rejecting fake tau candidates from jets, leaving the precise working point to be optimized in the context of a specific analysis. The cuts on the discriminating variable and p_T of the tau candidates were optimized with respect to a simple $S/\sqrt{S+B}$ performance measure, in signal events of the lh-channel to have only one true hadronically decaying τ. The optimization procedure for the hadronic-τ identification is as follows [76]. For each event, the τ-candidate with the highest value of the discriminant variable is selected. Jets overlapping with this candidate are removed. After cuts are applied, $S/\sqrt{S+B}$ is calculated for each cut on the likelihood variable, in steps of 1. Then, the optimal working point is obtained by finding the cut on the discriminant that maximizes $S/\sqrt{S+B}$. This procedure yields a higher probability to find the true hadronically decaying τ rather than choosing the candidate with the highest transverse momentum.

The background sample included Z+jets, W+jets, and $t\bar{t}$+jets, which provide a representative mixture of real and fake taus. Our modeling of the jet fragmentation indicates that quark-initiated jets are more collimated and can have a 6-8 times higher fake rate than gluon-initiated jets. The relative abundance of real and fake tau candidates depends on the kinematic requirements imposed on the sample by the analysis cuts, thus the optimization should be performed after the final kinematic requirements. However, the limited size of Monte Carlo samples requires that only a subset of the criteria used in the final event selection are applied during the optimization. Several subsets of the final event selection criterion were evaluated. An efficiency between roughly 40-50% is obtained for a likelihood ratio between three and five, leading to a rejection of about 80-300. Figure 5.2 illustrates the signal significance $S/\sqrt{S+B}$ as a function of the LLH cut (left) and as a function of LLH and $\mathrm{p_T}$ cut on the hadronic-τ (right). We find that variation inside the working point intervals leads to minimal changes in $S/\sqrt{S+B}$. Furthermore the optimal cut on the discriminant variable is largely independent of a cut on the $\mathrm{p_T}$ of τ-candidates. After the optimization, log-likelihood ratio was required to be greater than 4, corresponding to an identification efficiency of about 50.0% and a fake jet selection efficiency of ∼1% for gluon-initiated jets and ∼2.5% for quark-initiated jets for a $\mathrm{p_T}$ around 45 GeV. The hadronic tau identification performance and the fake-jet tagging rate as a function of $\mathrm{p_T}$ are shown in Figure 5.3 (left) and (right) respectively.

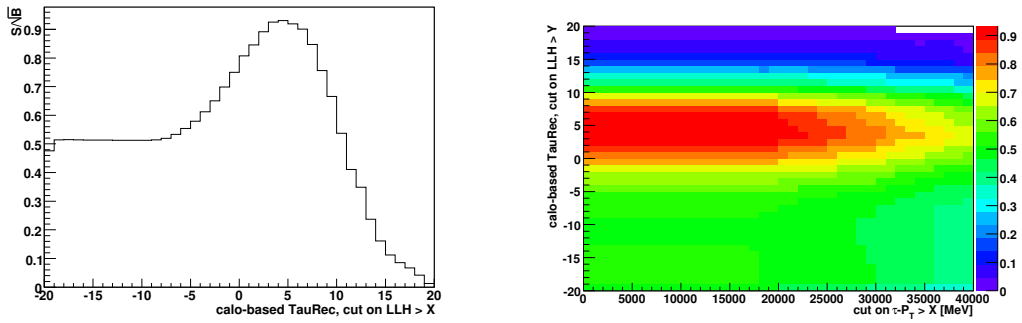

Figure 5.2: *Signal significance $S/\sqrt{S+B}$ as a function of the LLH cut (left) and as a function of LLH and $\mathrm{p_T}$ cut on the hadronic-τ (right).*

Moreover, an electron-veto was used to reject tau candidates which arise from electrons that have failed the electron identification. This electron-veto was performed by requiring the tau candidate to have at least 0.2% of its energy in the first sampling of the hadronic calorimeter and that the ratio of high-threshold (HT) to low-threshold (LT) hits in the transition radiation

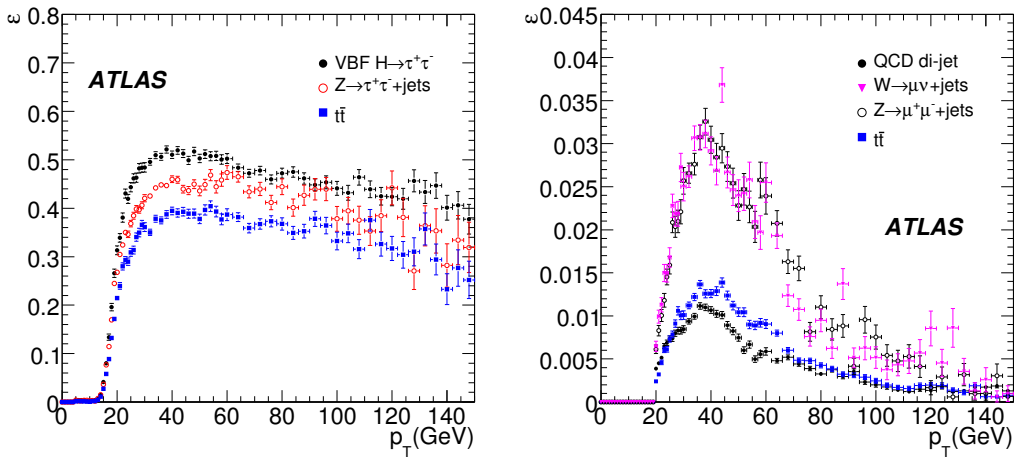

Figure 5.3: *Identification efficiency of the hadronic tau (left) and the jet-fake rejection efficiency (right) as a function of* p_T*, respectively.*

tracker (TRT) be less than 20% in the range $|\eta_\tau| < 1.7$. This electron-veto procedure suppresses the electron fake rate by 82.5% while retaining 90% of the hadronic tau candidates selected without the veto. Further requests on the track multiplicity and on the charge of the τ candidate are applied. The selection criteria for the hadronic tau identification are summarized in Table 5.5.

Hadronic tau identification
Tau ID: Calorimeter-seeded
$p_T \geq 30$ GeV
Track multiplicity : 1 or 3 tracks
$\lvert charge \rvert = 1$
Log Likelihood Ratio ≥ 4
Electron Veto:
minimum TRT $HT/LT \leq 0.2$ if $\lvert\eta_\tau\rvert \leq 1.7$ and $LT \geq 10$
$E_T^{HAD}/p_T \geq 0.002$ in matched electron object

Table 5.5: *Selection criteria for the hadronic tau identification from the calorimeter-seeded reconstruction algorithm.*

5.3.3 Forward Jet Tagging

The jet activity of the vector boson fusion process is unique in several ways, providing many handles to suppress backgrounds and isolate a sample of signal events with high purity. The most important features of the VBF process are the presence of two $high-\mathrm{p_T}$ quark-initiated "tagging jets", which tend to be relatively forward and well separated in rapidity. Furthermore, due to color coherence in this electroweak process, additional QCD radiation between the tagging jets tends to be suppressed and motivates a Central Jet Veto (CJV) [77] that is described in the next section.

Figure 5.4 shows the η spectra of the highest and second highest $\mathrm{p_T}$ jets in the signal and various background samples. Because the VBF jets can be very forward, the jet finding efficiency in this region is important in the analysis. Furthermore, the forward calorimeters ($3.1 \leq |\eta| \leq 4.9$) do not have a projective geometry, which leads to different challenges for jet reconstruction. ATLAS currently provides jets based on two algorithms, a seeded cone algorithm with split-merge and a k_T algorithm [78]. Each algorithm has two sets of parameters (the cone size and the k_T distance parameter), applied to two different input representations of the energy deposits in the calorimeter, towers merged to avoid negative energy fluctuations from electronic noise and clusters based on the ATLAS TopoCluster algorithm [54]. These different jet algorithms and the different calorimeter pre-clustering result in different performances for jets, especially at low $\mathrm{p_T}$ and high $|\eta|$.

Jet identification efficiency and purity are defined to give a quantitative measure of the jet identification [79]. The efficiency and purity were calculated with respect to generator-level jets obtained by running the same jet algorithm on the stable interacting particles after hadronization and before GEANT simulation. To ensure that only hadronic jets are considered, only dimuon events have been used, with both taus decaying into muon and neutrinos, or Z/W bosons directly decaying into muons. This avoids any bias in the jet reconstruction due to the presence of electrons. A reconstructed jet is considered to be matched if the corresponding generator-level jet is within a radius $\Delta\mathrm{R} \leq 0.15$ for jets with a cone size of 0.4. The matching cone size was chosen to avoid that a single generator-level jet were matched to more than one reconstructed jet; with the given parameters this effect is at the order of 10^{-3}.

The jet reconstruction efficiency in different $|\eta|$ regions and for two different clustering algorithms is shown in Figure 5.5 as a function of the generator-level jet $\mathrm{p_T}$ and η. The reconstruction efficiency rises over 95% for jets with $\mathrm{p_T}$ above 50 GeV. On the other hand, the efficiency drops at $|\eta| \sim 1.5$ and $|\eta| \sim 3.2$ for jets in the range of 20-30 GeV of $\mathrm{p_T}$. This

drop in efficiency is due to the crack region in the calorimeter or to the large amount of dead material in the corresponding η region. The jets based on calorimeter towers show a drop in efficiencies in the forward region, while the jets based on TopoClusters do not show this loss of efficiency [80, 81]. The splitting of cell signals to fill towers, which is predominant in the forward region where the cells are larger than the tower bin size, generates less seeds than the more integrating cell clustering, where cell signals are summed up rather than split, and thus create more likely signal objects above the seed threshold [78]. For this reason, jets based on TopoClusters have been chosen for this analysis.

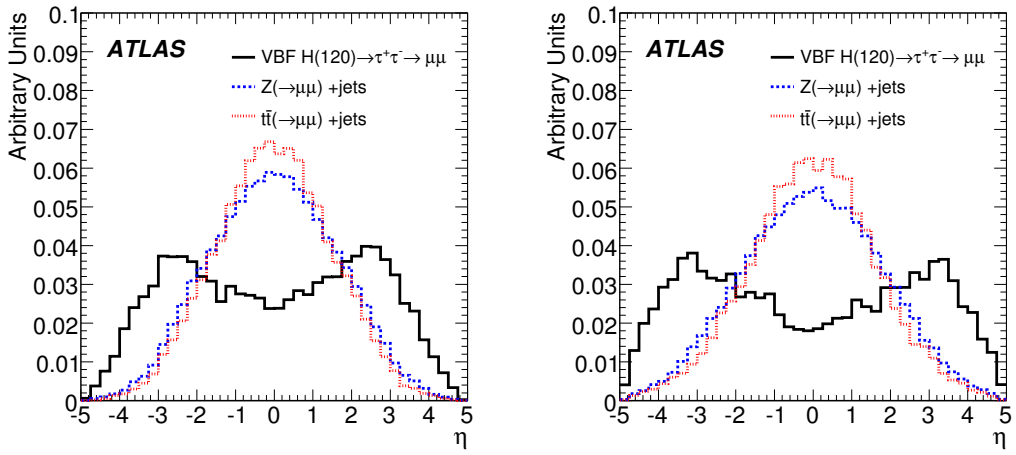

Figure 5.4: *Pseudorapidity of the highest* p_T *(left) and the second highest* p_T *(right) jets for the Cone jet algorithm based on TopoClusters with* $R = 0.4$ *in VBF* $H \to \tau\tau \to \mu\mu$ *(*m_H*=120 GeV) and background events. Only* p_T *cuts were applied to jets. Solid (black) histogram is for signal, dashed (red) histogram is for* $t\bar{t} \to WW \to (\mu\mu)$*, and dotted (blue) histogram is for* $Z \to \mu\mu$*+n jets.*

Correctly identifying the quark-initiated tagging jets from the VBF process is very important for the measurement of Higgs boson spin and CP properties and for making precise comparison with theoretical calculations [82]. Typically, the tagging jets are found in opposite hemispheres, but there are two approaches to incorporate this requirement in the analysis. One option is to define the tagging jets as the two highest p_T jets in the event, and reject the event from the signal candidates if they are in the same hemisphere (e.g. require $\eta_{j1} \times \eta_{j2} \leq 0$). A second option is to define the first tagging jet to be the highest p_T jet in the event and the second tagging jet to be the highest p_T jet in the opposite hemisphere. In this second approach it is not required that the second tagging jet is the second highest p_T jet in the event. These

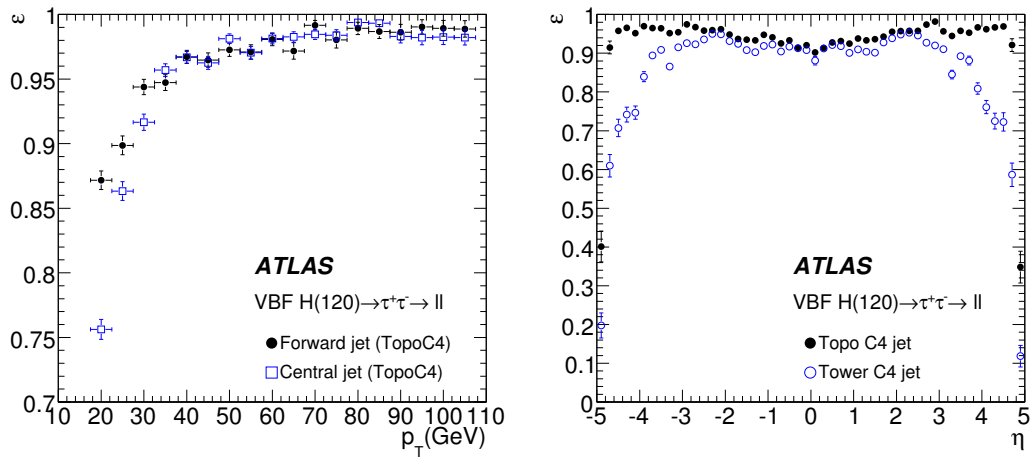

Figure 5.5: *Jet reconstruction efficiency for the Cone jet algorithm with $R = 0.4$ as a function of the generator-level jet* p_T *(left) and η (right) for Tower- and TopoCluster-based jets.*

two strategies were compared, and it was found that the first method more reliably matched the quark-initiated tagging jets from the hard process.

The generator-level jets match the hard-scattered quarks nearly 100% of the time above a p_T threshold of about 10 GeV. To estimate the purity of the tagging jets, the efficiency is calculated with respect to the generator-level jets. The reconstructed tag jets have a high purity over the entire p_T and η range and do not show a strong dependence on the jet algorithms.

Integrated efficiencies and purities for jets with $p_T \geq 20$ GeV indicate that the TopoCluster-based algorithm has better performance for this analysis. Because additional jets often lie in the central detector region, where we wish to employ a central jet veto, jets with smaller cones are favored for the selection to reduce the number of fake jets. Furthermore, calorimeter noise (including effects from pileup of minimum-bias events) increases with jet cone radius. Thus, we use the cone jet algorithm with $R = 0.4$ running on TopoClusters as the primary jet algorithm in this analysis.

The p_T cuts on the tagging jets are effective to reduce several backgrounds and Figure 5.4 shows that the pseudorapidity distributions are substantially different. Instead of relying directly on the pseudorapidity of the tagging jets, Figure 5.6 shows that the pseudorapidity gap (left) and invariant mass of the two tagging jets (right) provide substantial background rejection.

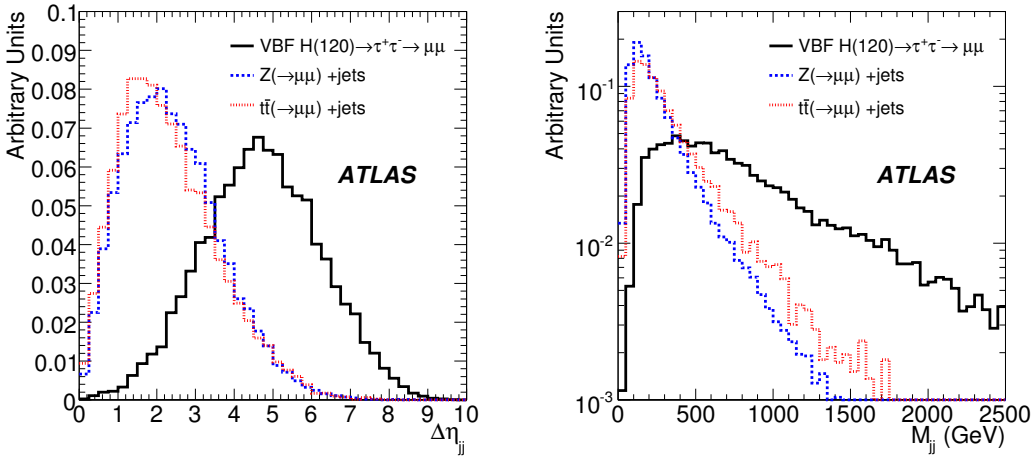

Figure 5.6: *Pseudorapidity gap between tag jets (left) and invariant-mass distributions of tag jets (right) in VBF $H \rightarrow \tau\tau \rightarrow \mu\mu$ events (m_H=120 GeV). A requirement $\eta_1 \times \eta_2 \leq 0$ is used in addition to the cuts on jet p_T. Solid (black) histogram is for signal, dashed (red) histogram is for $t\bar{t} \rightarrow WW \rightarrow (\mu\mu)$, and dotted (blue) histogram is for $Z \rightarrow \mu\mu$+n jets.*

5.3.4 Central Jet Veto

As mentioned above, the color coherence in the VBF Higgs boson production leads to a suppression of QCD radiation between the tagging jets. This color coherence is also found in the electroweak Z+jets background. In contrast, most of the other backgrounds have a much larger probability for additional QCD radiation in the central region. This is the physical motivation for a central jet veto (CJV). Figure 5.7 shows the jet multiplicity distribution for the signal and backgrounds after requiring two tagged jets (with $p_T \geq$ 20 GeV) in opposite hemispheres. The fraction of signal events with three or more jets is small. The experimental challenge for the CJV is to provide a cut that is robust against additional minimum bias events (in-time pileup events). The optimization for the central jet veto has been studied in terms of p_T and η. The probability to have at least one reconstructed jet with $p_T \geq 20$ GeV within $|\eta| \leq 3.2$ is 1.6% from a single minimum-bias event. In Figure 5.8, the trade-off of background rejection versus signal efficiency is shown for different p_T threshold on the third highest p_T jet (markers indicate thresholds of 20 and 30 GeV). A veto based on a fixed η-window was compared to a dynamic η-window defined by the η of the two tagging jets. We maintain the previous central jet veto requirement: no jets in $|\eta| \leq 3.2$ with $p_T \geq 20$ GeV.

The CJV poses significant theoretical challenges as well. At the parton-level, the CJV efficiency is expected to be known quite precisely with little

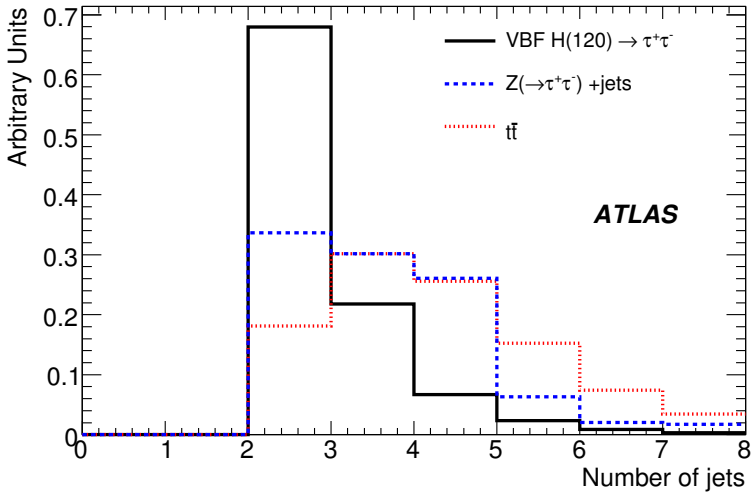

Figure 5.7: *Jet multiplicity distribution for the signal, Z+jets, and $t\bar{t}$ background after requiring the cuts up to the N jets ≥ 2 level in the list of cuts.*

theoretical uncertainty. However, the current tools that allow for the full parton-shower and hadronization (a prerequisite for any analysis based on a GEANT-based detector simulation) show significant uncertainties. We have observed significant differences between the central jet activity in signal events generated with PYTHIA and those generated with HERWIG [83]. Knowledge of the uncertainty on the CJV is needed for setting limits on the Higgs boson cross-section and for making coupling measurements; however, it is not needed directly in establishing a deviation from the background-only expectation.

In future studies we will also include a veto procedure using track information; in particular using vertexing information to reduce the impact of jets from in-time pileup. Furthermore, a track-based veto and the use of timing information in the calorimeter will also be studied to reduce the impact of out-of-time pileup.

5.3.5 Missing Transverse Energy

Significant missing transverse energy $\mathrm{E}_{\mathrm{T}}^{miss}$ is present in $H \rightarrow \tau^{+}\tau^{-}$ events because neutrinos are always associated with the τ decays. As discussed in section 4.2, the performance of the $\mathrm{E}_{\mathrm{T}}^{miss}$ algorithm plays a crucial role in this analysis because $\mathrm{E}_{\mathrm{T}}^{miss}$ is used in the mass reconstruction of the tau pair. Ultimately, the $\mathrm{E}_{\mathrm{T}}^{miss}$ resolution is what limits the $m_{\tau\tau}$ resolution. Furthermore, the absolute scale of the $\mathrm{E}_{\mathrm{T}}^{miss}$ must be well calibrated to correctly

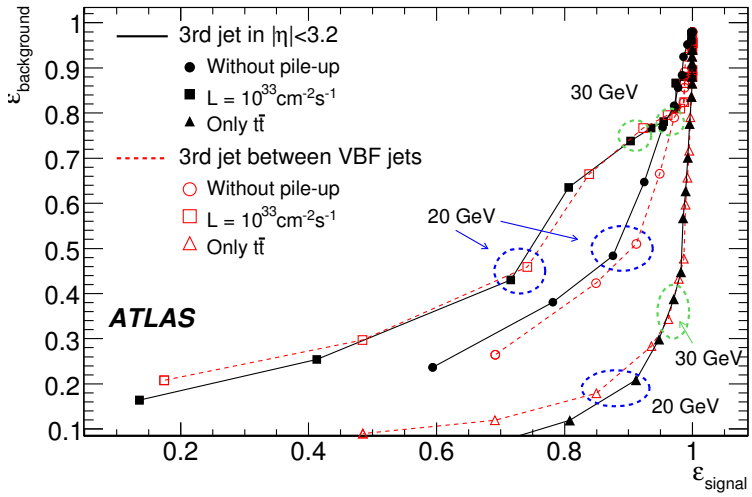

Figure 5.8: *Background rejection versus signal sensitivity for the central jet veto with and without pileup. Also shown is the case for $t\bar{t}$-only background.*

reconstruct the Higgs boson mass. By requiring a large E_T^{miss}, it is possible to improve the $m_{\tau\tau}$ resolution and reject many backgrounds that do not contain neutrinos (e.g. QCD dijet background). In Figure 5.9 the E_T^{miss} distribution is shown for signal events and for three QCD dijet samples. These dijets samples have been generated in three different p_T bins: 17-35 GeV (J1 sample), 35-70 GeV (J2 sample), 70-140 GeV (J3 sample). For the analysis of the hh-channel we require $E_T^{miss} \geq 40$ GeV.

5.3.6 Mass Reconstruction

Although there are several neutrinos in the event, it is possible to reconstruct the $\tau^+\tau^-$ invariant mass by making the approximation that the decay products of the τ are collinear with the τ in the laboratory frame. This is a good approximation since $m_H/2 \gg m_\tau$, thus the τs are highly boosted. This leaves two unknown quantities and two equations: the fraction of each τ's momentum carried away by neutrinos and the constraints from the two components of E_T^{miss}. Consider the hh-channel and let h_1 and h_2 represent the momentum vectors for the hadronic visible decay products. By neglecting the τ rest mass and imposing the collinear approximation, we can write

$$m_{\tau\tau} = \sqrt{2(E_{h_2} + E_{\nu h_2})(E_{h_1} + E_{\nu h_1})(1 - \cos\theta_{h_1 h_2})}\ , \tag{5.1}$$

where E_h and $E_{\nu h}$ are the energies of the visible decay products and of the neutrino respectively, and $\theta_{h_1 h_2}$ is the angle between the two τs. By

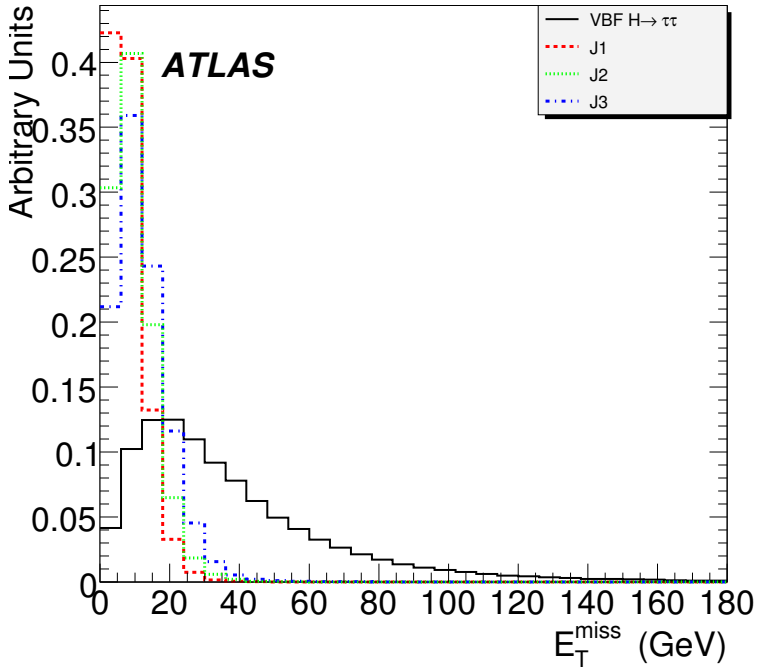

Figure 5.9: E_T^{miss} *distribution for the signal* $VBF\ H \rightarrow \tau\tau$ *(solid black line) and for three QCD dijet samples generated in different* p_T *bins: 17-35 GeV for J1 sample (dashed red line), 35-70 GeV for J2 sample (dotted green line), 70-140 GeV for J3 sample (dot-dashed blue line).*

introducing new variables x_{h_2} and x_{h_1}, representing the fraction of the τ's momentum carried away by the visible decay products, we can re-write the invariant mass as

$$m_{\tau\tau} = \frac{m_{h_1 h_2}}{\sqrt{x_{h_1} x_{h_2}}} \qquad \text{for } x_{h_1,h_2} \geq 0 \ . \tag{5.2}$$

One can easily solve for the x_h variables by requiring that the vector sum of the neutrinos coincides with the two measured components of $\mathrm{E}_{\mathrm{T}}^{miss}$:

$$x_{h_2} = \frac{E_{h_2}}{E_{h_2} + E_{\nu h_2}} = \frac{h_{2_x} h_{1_y} - h_{2_y} h_{1_x}}{h_{2_x} h_{1_y} + E_x^{miss} h_{1_y} - h_{2_y} h_{1_x} - E_y^{miss} h_{1_x}} = \frac{N}{D_{h_2}} \ , \tag{5.3}$$

and

$$x_{h_1} = \frac{E_{h_1}}{E_{h_1} + E_{\nu h_1}} = \frac{h_{2_x} h_{1_y} - h_{2_y} h_{1_x}}{h_{2_x} h_{1_y} - E_x^{miss} h_{2_y} - h_{2_y} h_{1_x} + E_y^{miss} h_{2_x}} = \frac{N}{D_{h_1}} \ , \tag{5.4}$$

where we have introduced the quantities N, D_{h_2}, and D_{h_1} for convenience. If the two τs are back-to-back, then these equations are linearly-dependent and one cannot solve for the x_hs. For this reason, indicating with ϕ the angle in the transverse plane, we require that $\cos\Delta\phi_{h_1 h_2} \geq -0.9$. Typically the Higgs boson has significant $\mathrm{p_T}$ due to the tagging jets. Events that come from the process $X \to \tau\tau$ with no other sources of missing energy should have $0 \leq x_\tau \leq 1$, though resolution effects in $\mathrm{E}_{\mathrm{T}}^{miss}$ may lead to unphysical solutions with either $x_\tau < 0$ or $x_\tau > 1$.

5.3.7 Total $\mathrm{p_T}$ and Di-tau Transverse Mass

Other two variables have been identified as potentially useful for the background rejection. Considering the two identified tau jets, the two forward tagging jets and the $\mathrm{E}_{\mathrm{T}}^{miss}$ of the event, the signal events appear more balanced in the transverse plane than background events with different topology. The total transverse momentum $\mathrm{p}_{\mathrm{T}}^{total}$ given below can be used to discriminate signal and background:

$$\mathrm{p}_{\mathrm{T}}^{total} = ||\vec{\mathrm{p}}_{\mathrm{T}}^{\,h1} + \vec{\mathrm{p}}_{\mathrm{T}}^{\,h2} + \vec{\mathrm{p}}_{\mathrm{T}}^{\,j1} + \vec{\mathrm{p}}_{\mathrm{T}}^{\,j2} + \vec{E}_T^{miss}|| \ . \tag{5.5}$$

In Figure 5.10 (left) the distribution of the total $\mathrm{p_T}$ is shown for the signal and for $t\bar{t}$ events.

Also variables related to the di-tau system mass can be used to reduce those backgrounds where there are one or two fake τs reconstructed in the

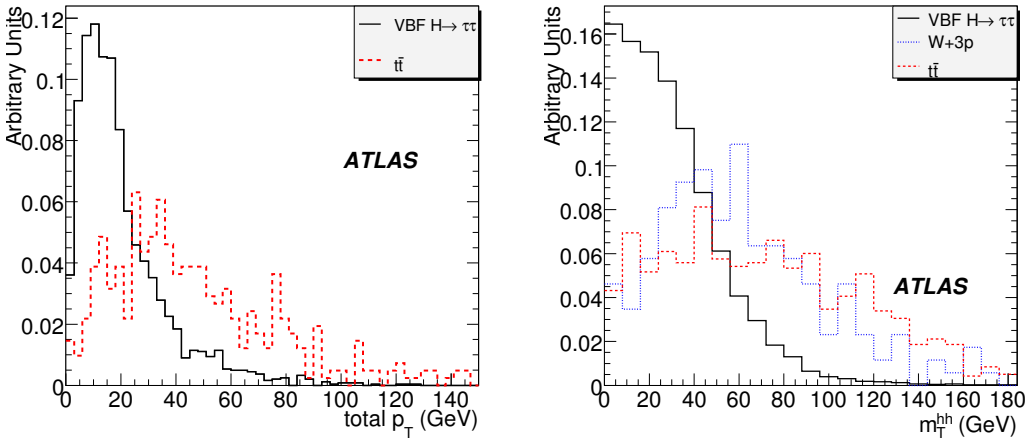

Figure 5.10: *Distribution of the total transverse momentum* $\mathrm{p_T^{total}}$ *for the signal* $VBF\ H \to \tau\tau$ *(solid black line) and for the* $t\bar{t}$ *background (dashed red line) (left).* m_T^{hh} *distribution for the signal* $VBF\ H \to \tau\tau$ *(solid black line), the* $t\bar{t}$ *(dashed red line) and* $W + 3jets$ *(dotted blue line) backgrounds (right)*

event. Figure 5.10 (right) shows the distribution of the di-tau transverse mass defined as:

$$m_T^{hh} \;=\; \sqrt{2\ \mathrm{p_T^{hh}}\ \mathrm{E_T^{miss}} \cdot (1 - \cos\Delta\phi)}\ , \tag{5.6}$$

where $\mathrm{p_T^{hh}}$ is the transverse momentum of the two hadronic tau system and $\Delta\phi$ represents the azimuthal angle between $\mathrm{p_T^{hh}}$ and $\vec{E}_T^{miss}$, for signal events and for background events from $t\bar{t}$ and $W + 3jets$.

5.3.8 Summary of Event Selection

The final event selection for hh-channel is summarized below.

- Trigger: a combination of the hadronic tau and missing E_T trigger `tau35i_xe40`.
- Hadronic taus: two identified hadronic taus are required with $\mathrm{p_T}$ above 35 GeV and 30 GeV and with opposite charge.
- Missing $\mathrm{E_T}$: $\mathrm{E_T^{miss}} \geq 40$ GeV.
- Collinear approximation: $0.2 \leq x_{h1,h2} \leq 1$, and $\cos\Delta\phi_{hh} \geq -0.9$.
- Di-tau transverse mass: in order to further suppress fake-τ candidates from W + jets and $t\bar{t}$ backgrounds, a cut $m_T^{hh} \leq 80$ GeV is required.

This variable has been identified as potentially useful for the analysis. The optimal value of this cut depends heavily on the relative amount of the W+jets, $t\bar{t}$ and QCD backgrounds, therefore, the requirement is kept fairly loose.

- Jet multiplicity: at least one jet with $p_T \geq 40$ GeV and at least one additional jet with $p_T \geq 20$ GeV.
- Total p_T: $p_T^{total} \leq 60$ GeV.
- Forward jets: the two tagging jets must be in opposite hemispheres $\eta_{j1} \times \eta_{j2} \leq 0$, and the two tau-jets must lie between them: $\min\{\eta_{j1}, \eta_{j2}\} \leq \eta_{\tau_1,\tau_2} \leq \max\{\eta_{j1}, \eta_{j2}\}$.
- Jet kinematics: $\Delta\eta_{jj} \geq 4$ and dijet mass $m_{jj} \geq 700$ GeV for two forward jets.
- Central jet veto: the event is rejected if there are any additional jets with $p_T \geq 20$ GeV with $|\eta| \leq 3.2$.
- Mass window: $m_H - 15$ GeV$\leq m_{\tau\tau} \leq m_H + 20$ GeV around the test mass m_H.

Table 5.6 summarizes the cross-section for signal events after each of the cuts described above.

Higgs Mass (GeV)	105	110	115	120	125	130	135
Cross section (fb)	394.7	372.0	341.8	309.1	266.8	225.4	180.1
Trigger τ & E_T^{miss}	12.4(2)	12.1(2)	12.0(2)	11.4(1)	10.4(2)	9.2(1)	7.9(1)
2 Hadronic τs	1.73(8)	1.80(8)	1.93(8)	1.83(4)	1.67(7)	1.52(5)	1.29(5)
$E_T^{miss} \geq 40$ GeV	1.34(7)	1.39(7)	1.50(7)	1.43(3)	1.32(6)	1.17(5)	0.99(4)
Collinear Approx.	0.91(6)	1.02(6)	1.13(6)	1.03(3)	1.00(5)	0.85(4)	0.72(3)
Di-tau m_T^{hh}	0.91(6)	1.02(6)	1.13(6)	1.03(3)	1.00(5)	0.85(4)	0.72(3)
N jets ≥ 2	0.77(5)	0.88(6)	0.94(5)	0.86(3)	0.84(5)	0.72(4)	0.61(3)
Total p_T	0.72(5)	0.84(5)	0.91(5)	0.83(3)	0.80(5)	0.69(4)	0.58(3)
Forward jet	0.62(5)	0.73(5)	0.75(5)	0.72(2)	0.68(4)	0.58(3)	0.50(3)
Jet kinematics	0.37(4)	0.43(4)	0.41(4)	0.45(2)	0.41(3)	0.36(3)	0.28(2)
Central jet veto	0.34(3)	0.38(4)	0.36(3)	0.39(2)	0.35(3)	0.32(3)	0.24(2)
Mass window	0.25(3)	0.35(4)	0.33(3)	0.34(2)	0.29(3)	0.27(2)	0.20(2)

Table 5.6: *Signal cross-sections for the hh-channel for various Higgs boson masses.*

Figure 5.11 shows the reconstructed mass after all the selection cuts for a Higgs boson with a mass of 120 GeV. The reconstructed mass peak is at

116 GeV and it is shifted of about 4 GeV with respect to the nominal value. This is due to the fact that the standard E_T^{miss} algorithm uses the standard jet calibration even for tau objects. A dedicated tau correction in the E_T^{miss} calculation can remove the bias [76].

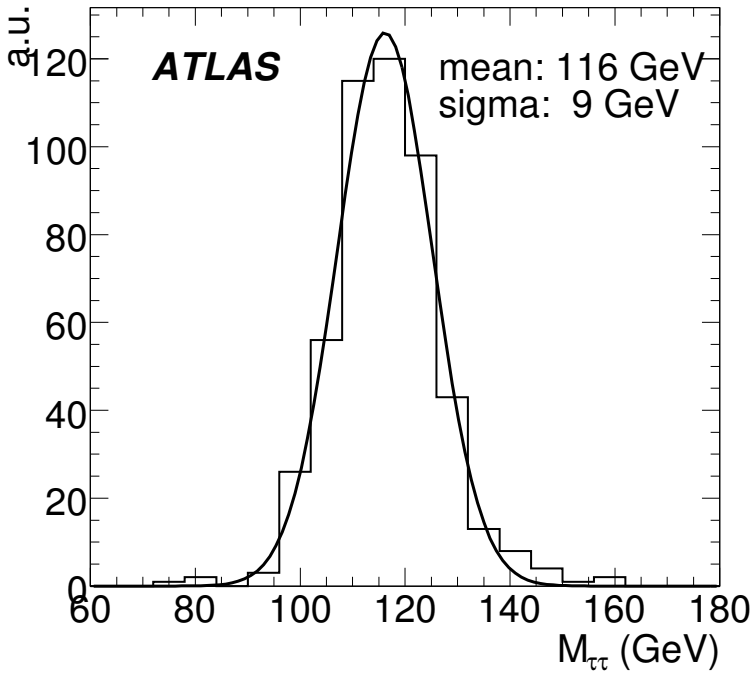

Figure 5.11: *Reconstructed mass of a 120 GeV Higgs boson with the collinear approximation method.*

A first evaluation of effect of pileup can be given on the signal events. The presence of pileup has three major effects on the analysis. First, additional $p-p$ interactions can produce hadronic activity in the central region which causes events to fail the central jet veto. Secondly, the presence of pileup generally degrades the E_T^{miss} resolution, which, in turn, reduces the efficiency of the collinear approximation cuts and degrades the $m_{\tau\tau}$ resolution and thus requires a broader mass window with a lower signal-to-background ratio for the same signal efficiency. Thirdly, pileup degrades the hadronic tau identification.

In the first column of the Table 5.7 the effect of the cut flow is shown for events simulated with the pileup at L=10^{33}cm^{-2}s^{-1}. The reconstructed τs are required to have a likelihood ratio bigger than four. It can be seen that a lower efficiency is obtained at the end of the cuts. This is due to the fact that both the likelihood function and the E_T^{miss} reconstruction algorithm have

not been tuned for the pileup. The distribution of the log likelihood ratio discriminant for the calorimeter-based hadronic tau identification algorithm is shifted to lower values for both real taus and jet fakes. By simply adjusting the cut on the log likelihood ratio to 0, the same signal efficiency can be maintained with approximately a 50% drop in jet rejection. This is shown in the second column in Table 5.7. By re-tuning the discriminant in the context of pileup, improved jet rejection is possible.

In Figure 5.12 the reconstructed mass peak for the pileup at 10^{33} is shown. The mass resolution is degraded from ∼9 to ∼ 12 GeV for $m_H = 120$ GeV due to the degradation of the $\mathrm{E}_{\mathrm{T}}^{miss}$ resolution.

5.4 Background Estimation

The major challenge in the prediction of background is related to the limited size of our Monte Carlo samples. Backgrounds from mis-identified leptons are difficult to estimate due to the large rejection factors of the identification algorithms. Even the irreducible $Z \rightarrow \tau\tau$ backgrounds are suppressed by several orders of magnitude due to the kinematic requirements. The $t\bar{t}$ background requires particularly large sample sizes because it is suppressed by both identification and kinematic requirements. In addition to the Z+jets, W+jets and $t\bar{t}$ backgrounds, the hh-channel also suffers from a background from the pure QCD multijet process.

In table 5.8 the summary of the selection cuts is reported, when attempting to run on full simulated samples applying cuts in correct sequence, with the mass window cut set with respect to the test Higgs mass of 120 GeV. It can be seen that with the available background statistics no event is left at the mass window cut. For QCD dijets in particular, there is no event left already at the early stages of the selection, hence the decision to move to ATLFAST samples for this background.

To predict the final background rates for $t\bar{t}$, $Z + jets$ and $W + jets$ it was thus necessary to use a cut factorization method. The prediction of the pure QCD multijet background from Monte Carlo is hopeless without factorizing the analysis even further. A sample of 80 million QCD dijet events (including $c\bar{c}$ and $b\bar{b}$ processes), with $17 < \mathrm{p_T} < 240$ GeV, were simulated with the fast detector simulation. The tau fake rate was parameterized from full simulation as a function of η and $\mathrm{p_T}$ and used to re-weight the dijet sample.

Mass (GeV)	120	120
	$(\tau - id\ LLH > 4)$	$(\tau - id\ LLH > 0)$
Cross section (fb)	309.1	309.1
Trigger tau & MET	12.4(2)	12.4(2)
2 Hadronic τs	1.44(5)	2.76(7)
$E_T^{miss} \geq 40$ GeV	1.07(5)	1.99(6)
Collinear Approx.	0.69(4)	1.26(5)
Di-tau Transverse mass	0.69(4)	1.26(5)
N jets ≥ 2	0.57(3)	1.05(4)
Total p_T	0.54(3)	1.01(4)
Forward jet	0.48(3)	0.87(4)
Jet kinematics	0.30(2)	0.54(3)
Central jet veto	0.24(2)	0.42(3)
Mass window	0.20(2)	0.34(3)

Table 5.7: *Signal efficiency with pileup at $L=10^{33} cm^{-2} s^{-1}$. In the first column the standard τ-identification criteria are applied; in the second column the value of the likelihood ratio function is lowered to zero to get the same overall efficiency obtained without pileup.*

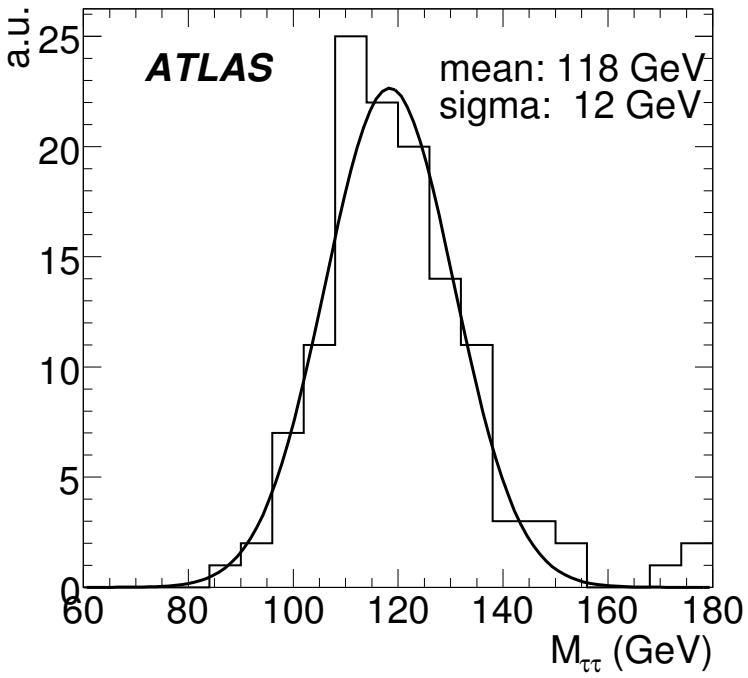

Figure 5.12: *Reconstructed mass of a 120 GeV Higgs boson with the collinear approximation method, in events with pileup for $L=10^{33} cm^{-2} s^{-1}$.*

	$Z \to \tau\tau$		$t\bar{t}$	$W \to \tau\nu$	QCDjets
	QCD	ELWK			
Cross section (fb)	40.3×10^3	1693	833×10^3	922×10^3	$19.10(2)\times10^{12}$
Trigger tau & MET	1756(15)	126(1)	78177(232)	39600(400)	8.740×10^7
2 Hadronic τs	161(4)	4.9(2)	373(16)	317(33)	0
$E_T^{miss}\geq$40 GeV	108(4)	3.7 (2)	335(15)	243(29)	0
Collinear Approx.	72(3)	2.3 (1)	43(5)	20(7)	0
Di-tau m_T^{hh}	72(3)	2.3(1)	39(5)	18(7)	0
N jets ≥ 2	58(2)	2.1(1)	38(5)	11(5)	0
Total p_T	49(2)	1.9(1)	24(4)	11(5)	0
Forward jet	21(1)	1.1(1)	10(3)	4(3)	0
Jet kinematics	1.1(3)	0.43(6)	1(1)	2(2)	0
Central jet veto	0.4(2)	0.36(6)	1(1)	2(2)	0
Mass window	0	0.03(1)	0	0	0

Table 5.8: *Signal and background cross sections for the hh channel as a function of the sequential cuts of the analysis. Numbers in brackets represent the errors on the cross sections.*

5.4.1 Cut Factorization Method

The analysis cuts can be divided into two categories. The first category is related to the tau decays from the Higgs boson candidate (trigger, hadronic tau ID, E_T^{miss}, the collinear approximation, and the di-tau transverse mass) and the rejection is dominated by detector effects. The second category of cuts is related to the tagging of jets (forward jets, jet separation, and dijet mass) and the rejection is dominated by the kinematic properties of the events. In our approach, the rejection of the jet-related cuts were calculated without any tau-related cuts. The dropped cuts are: trigger, E_T^{miss}, collinear approximation, transverse mass. Then the background cross sections are normalized after the collinear approximation. The goodness of the method can be checked comparing the cross section for factorized and non-factorized cut flow after the transverse mass cut. Examples of checks of the factorization procedure are shown for $t\bar{t}$ and $Z+2p$ in Table 5.9 and Table 5.10 respectively.

The agreement between the factorization and the direct cut flow can be assessed comparing the cross section for sequential and factorized cuts. For the $t\bar{t}$ the two cross sections are well consistent within the errors. For the $Z + 2p$ there is a 40% difference in the cross section of the two methods and the consistency is recovered at the level of jet kinematics cut. Similar checks have been done also for the other samples and with the available statistics the cross section from the factorization procedure is in agreement with the direct cut flow within the errors. However, using the factorization procedure for $t\bar{t}$ and $W + jets$ no event remains at the mass window. Conservatively,

Process (fb)	$t\bar{t}$	Factorized cuts	Applied
Cross section	833×10^3	833×10^3	
Trigger tau & MET	78177(232)	833×10^3	
2 Hadronic τs	373(16)	823(24)	
$E_T^{miss} \geq 40$ GeV	335(15)	823(24)	
Collinear Approx.	43(5)	823(24)	
Transverse mass	39(5)	823(24) (100 %)	39(5)
N jets ≥ 2	38(5)	732(22) (89±1 %)	34(5)
Total p_T	24(5)	509(19) (62±1 %)	24(5)
Forward jet	10(3)	190(11) (23±1 %)	9(2)
Jet kinematics	0.7(7)	12(3) (1.5±0.3 %)	0.6(2)
Central jet veto	0.7(7)	3(2) (0.4±0.2 %)	0.16(9)
Mass window	0	0.64(0.64) (0.08±0.08 %)	0.03(3)*

Table 5.9: *Comparison between sequential cuts and factorized cuts for the $t\bar{t}$ background. * No event in the mass window, thus one event is considered.*

Process (fb)	Z+2p	Factorized cuts	Applied
Cross section after filter	15240	15240	
Trigger tau & MET	694(11)	15240	
2 Hadronic τs	63(3)	567(10)	
$E_T^{miss} \geq 40$ GeV	42(3)	567(10)	
Collinear Approx.	29(2)	567(10)	
Transverse mass	28(2)	567(10) (100%)	28(2)
N jets ≥ 2	24(2)	335(7) (59±1%)	17(2)
Total p_T	24(2)	332(7)(59±1%)	17(1)
Forward jet	10(1)	127(5)(22±1%)	6.4(7)
Jet kinematics	0.5(3)	9(1)(1.7±0.2%)	0.5(1)
Central jet veto	0.3(2)	8(1)(1.4±0.2%)	0.4(1)
Mass window	0	1.1(4)(0.17±0.07 %)	0.05(2)

Table 5.10: *Comparison between sequential cuts and factorized cuts for the $Z+2p$ background.*

in this case one event left has been considered to estimate the final cross section.
The analysis of the ll and lh-channel [76], that also uses ATLFAST samples to estimate the rejection factors for $t\bar{t}$, shows that a factor 10 of rejection is obtained with the mass window cut. To estimate this rejection factor in our case, the likelihood value for τ-identification (see Table 5.5) has been weakened to zero and a factor of the order of 10 has been obtained also for the hh-channel. Similarly for the $W + jets$, about a factor 10 in rejection is obtained at the mass window cut by lowering the likelihood value for the τ-identification. A similar factor is found for ll and lh-channels using a factorization procedure. As a conclusion our choice to consider one event left for these backgrounds results in an upper limit estimation from their contribution.

5.4.2 Characteristics of ATLFAST

To understand the systematics originating from the use of ATLFAST samples in the weighting procedures used for the QCD background estimate, it is important to compare the reconstruction efficiency for jets and the $\mathrm{E}_{\mathrm{T}}^{miss}$ measurement for fast and full simulation: in our selection we require four jets and a $\mathrm{E}_{\mathrm{T}}^{miss}$ above 40 GeV.

The program ATLFAST [72] provides a fast simulation of the ATLAS detector response. The fast simulation allows to simulate large samples of background processes to study the background rejection, leaving in any case the critical performances to be addressed with a full simulation approach. To keep the CPU time per event at a reasonable level, no detailed simulations of any interactions of the particles in the detector media are performed. These interactions are implicitly taken into account by using a parametrized detector response [84].

Stable particles from the event generators are traced through the magnetic field. The primary vertex is assumed to be at the center of the detector which coincides with the center of the coordinate system. Assuming a perfect homogeneous magnetic field inside the central tracking volume, the impact point on the calorimeter surface is calculated. To calculate this point no interactions of the particle with the detector media are taken into account.

The energies of the electrons, photons and hadrons are deposited in a calorimeter cell map. The response of the calorimeter is assumed to be one and uniform over the full detector region. No smearing, i.e. no resolution function, is applied at this level. The energy of the particle is entirely deposited in the hit calorimeter cell, assuming a granularity of the calorimeter cell map of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ up to $|\eta| < 3.2$ and $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ for

$3.2 < |\eta| < 5.0$. No lateral nor longitudinal shower development is simulated. This implies that also the longitudinal fine structure of the calorimeters is not taken into account. There is no separation between the electromagnetic and the hadronic calorimeter compartment and no hits or track coordinates are simulated in the inner tracking detectors or in the muon chambers.

Based on the map of deposited energies a cluster reconstruction algorithm is carried out. A simple cone algorithm with a cone size of $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.4$ is used. The cone is initiated by seed cells with energies above 1.5 GeV. The cluster energy must pass a threshold and then the clusters are classified as electrons, photons, taus or jets.

The reconstruction of the physics objects in ATLFAST relies to a large extent on the Monte Carlo truth information. The cluster classification is done matching the true particles to the calorimetric clusters in (η, ϕ) space. First clusters are assigned to true electrons and photons. The remaining ones are considered as jets. Their energies are smeared according to the expected jet resolution separately in the barrel and in the forward region. After smearing, the non-isolated muon four vectors which lie in a cone with $\Delta R = 0.4$ from the jet are added to the jet itself. If the transverse energy of the candidate jet exceeds a 10 GeV threshold, it is accepted as jet.

To compare the jet reconstruction efficiency, the default E_T seed values for both the ATLFAST clustering and for the full reconstruction from calorimetric towers (E_T = 1.5 and 2 GeV respectively) are used. In order to compare jets of the same size, the same reconstruction cone has been used ($\Delta R = 0.4$). The matching criterion is $\Delta R = 0.2$ both for the fast and the full simulation. The results are shown in Figure 5.13. In general, the jet reconstruction efficiency is overestimated by ATLFAST with respect to the full simulation/reconstruction at low p_T . In the central region the two curves become consistent for p_T values above 40 GeV. The region $2.5 < |\eta| < 3.5$ shows a much worse efficiency for full simulation, which is linked to the modeling of the crack between the hadronic endcap and the forward calorimeters. The higher efficiency can increase the number of fast simulated events that pass the forward jets cuts.

The missing transverse energy is calculated from the reconstructed objects and from remaining clusters not associated to jets within the pseudorapidity range $|\eta| < 5$. In addition, cells not associated to clusters are included in the E_T^{miss} calculation. The components of the true missing transverse energy are calculated in ATLFAST from all neutrinos and from the non-visible muons. Muons are considered to be non-visible if they are outside the acceptance of the muon spectrometer ($|\eta|$ >2.7) or if their transverse momentum value is below the reconstruction threshold in the fiducial region (p_T <6 GeV and $|\eta|$ <2.7).

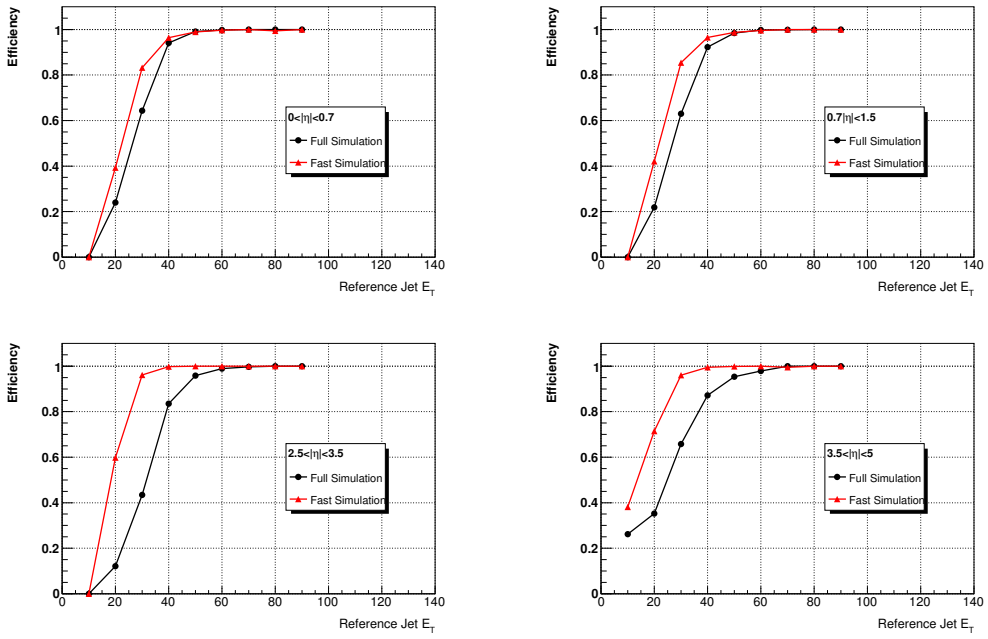

Figure 5.13: *The reconstruction efficiency as a function of the transverse momentum p_T^{MC} of the particle jet for ATLFAST (red) and the full reconstruction (black) for four different pseudorapidity regions.*

In ATLFAST, ΣE_T is calculated from the transverse energies of all objects used to reconstruct E_T^{miss} except muons. In the full simulation, ΣE_T is calculated from the transverse energies of the calorimeter cells used for the E_T^{miss} calculation (cells surviving the noise cut or inside TopoClusters). In the fast simulation a slightly better E_T^{miss} resolution is found. In addition, the ΣE_T calculation is in better agreement with the true value. This degradation can be explained by an increased noise in the calorimeter. In addition, the noise suppression and the calibration procedures applied are different. In Figure 5.14 the E_T^{miss} distributions obtained with fast (red dashed line) and full (black solid line) simulation are shown for two QCD dijet samples. These dijets samples have been generated in two different p_T bins: 35-70 GeV (left), 70-140 GeV (right). It can be seen that the distributions for fast an

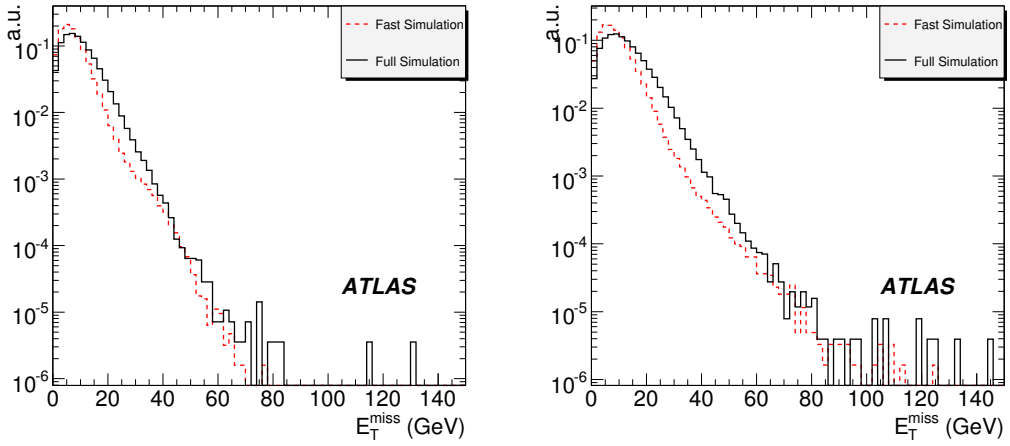

Figure 5.14: E_T^{miss} *distribution obtained with fast (red dashed line) and full (black solid line) simulation for two QCD dijet samples generated in different* p_T *bins: 35-70 GeV (left), 70-140 GeV (right).*

full simulation have a different shape. Even if the analysis requires a high E_T^{miss}, more events populate the tail of the full simulation distributions.

From the comparison between fast and full simulation we understand that there are significant differences between the two predictions. In principle, this discrepancy should be taken into account to calculate additional systematic error when using in particular ATLFAST, as we did for QCD background.

However, since at physics level already there are large uncertainties in QCD simulation, we think that final credible prediction for this background, physics plus detector effects, will be achieved only using the data themselves, as suggested in the section 5.6.

5.4.3 QCD weighting

The tau efficiency/rejection factors obtained by full simulation as a function of $\eta - p_T$ are used to weight the QCD dijet samples on event by event base. Figure 5.15 shows the rejection as a function of the jet p_T for two likelihood functions implemented in the tau offline reconstruction. They are different for the calorimetric seed used in the τ reconstruction: the red open squares (LLH2004) refer to the standard offline algorithm that uses the sliding window clustering algorithm (see section 4.1.2); the black full circles (LLH2007) refer to a new algorithm that uses TopoJets. Both likelihood functions are required to be greater than four. Values from LLH2004, the same one used for the τ-identification for the signal and background fully simulated events, have been considered to extract the weights. We expect from this alone a rejection factor against QCD jets ranging from $10^{-4} - 10^{-5}$. For every event, each jet configuration with two jets in the central region ($|\eta| < 2.5$) and with p_T above 35 and 30 GeV is weighted and then the cut sequence is applied. No trigger information is available for fast simulated events.

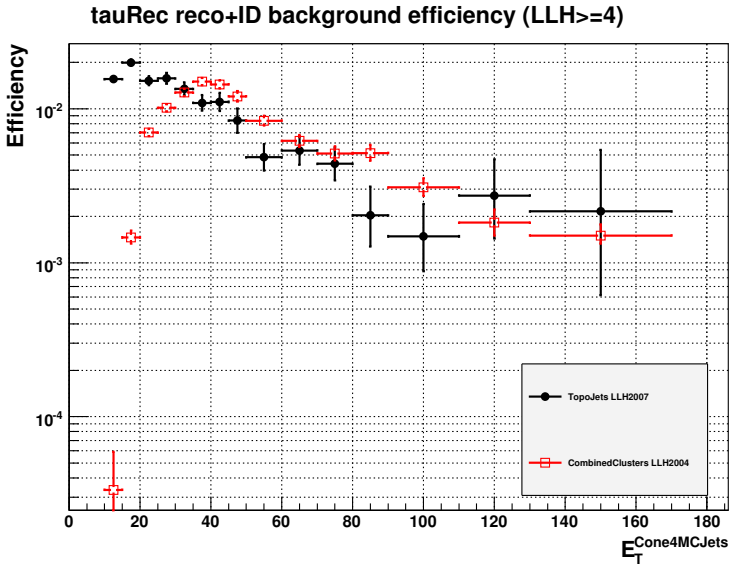

Figure 5.15: *Rejection as a function of the jet* E_T *for QCD jets, for two different* τ *reconstruction algorithms. Red open squares (LLH2004) refer to the standard offline algorithm that uses the sliding window clustering algorithm; black full circles (LLH2007) refer to a new algorithm that uses TopoJets.*

In the QCD dijet samples with $35 < p_T < 280$ GeV, starting from about 52 millions events, 65 events arrive at the central jet veto step, and for samples with $70 < p_T < 140$ GeV (13 millions events) six events survive at

the mass window cut. For jets with $35 < p_T < 70$ and $140 < p_T < 280$ GeV, the efficiency of the last cut is estimated applying the factorization method explained above. This efficiency is consistent with that one observed on the other QCD dijet sample with more available statistics. For the QCD dijet sample with $17 < p_T < 35$ GeV the factorization method has been used to provide final numbers at mass window.

No charge cut is applied to this background because in fast simulation this information is missing, while it is applied to all other backgrounds and signal, as they are analyzed using the full simulation samples. Therefore we expect an additional reduction of a factor two on the final estimated number in the mass window. Figure 5.16 shows the charge product of two reconstructed τs with one or three associated tracks, $LLH > 4$ and $p_T > 35$, 30 GeV in fully simulated QCD dijet events. It can be seen that a factor of 2 in rejection is obtained requiring the charge product of the two τs equal to -1.

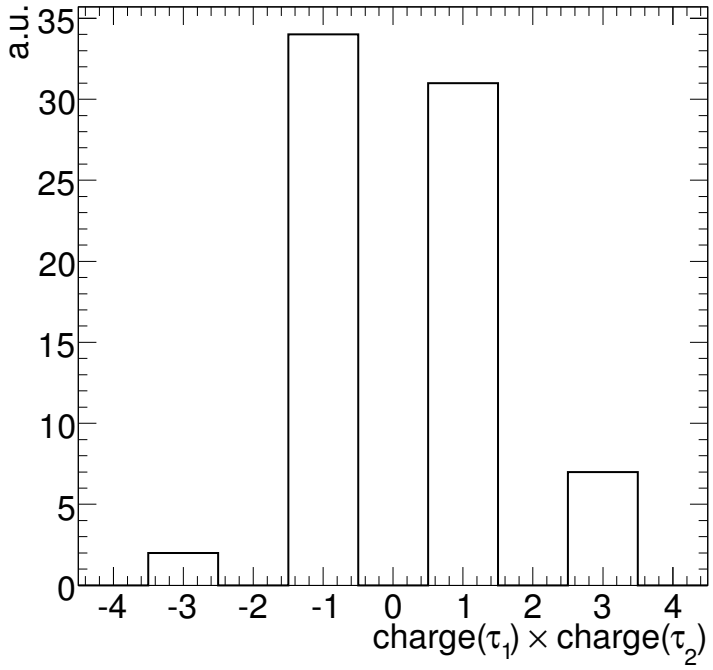

Figure 5.16: *Charge product of two reconstructed τs with one or three associated tracks, $LLH > 4$ and $p_T > 35$, 30 GeV in fully simulated QCD dijet events with $70 < p_T < 140$ GeV.*

5.4.4 Summary of the background prediction

In Table 5.11 we report the estimate of the background contribution.

	$Z \to \tau^+\tau^- + jets(\geq 1)$		$t\bar{t}$	$W \to \tau\nu +$	QCD dijet
	QCD	ELWK		$jets(\geq 1)$	
Cross section (fb)	40.3×10^3	1693	833×10^3	922×10^3	$19.1\ 10^{12}$
Trigger τ & E_T^{miss}	1756(15)	126(1)	78177(232)	39600(400)	
2 Hadronic τs	161(4)	4.9(2)	373(16)	317(33)	$2.756(3)\ 10^{6}$*
$E_T^{miss} \geq 40$ GeV	108(4)	3.7 (2)	335(15)	243(29)	$0.97(3)\ 10^{3}$*
Collinear Approx.	72(3)	2.3 (1)	43(5)	20(7)	$1.7(2)\ 10^{2}$*
Di-tau m_T	72(3)	2.3(1)	39(5)	18(7)	$1.6(2)\ 10^{2}$*
N jets ≥ 2	46(2)*	2.1(1)	34(5)*	8(3)*	$0.86(4)\ 10^{2}$*
Total p_T	40(2)*	1.9(1)	24(4)*	8(3)*	$0.75(3)\ 10^{2}$*
Forward jet	17(1)*	1.1(1)	9(2)*	3(1)*	4.5(5)*
Jet kinematics	1.4(1)*	0.43(6)	0.6(2)*	0.5(4)*	1.5(5)*
Central jet veto	0.7(1)*	0.36(6)	0.16(9)*	0.3(3)*	0.8(2)*
Mass window	0.08(3)*	0.03(1)	0.03(3)*	0.1(1)*	0.2(1)*

Table 5.11: *Summary of backgrounds for the hh-channel. An asterisk is used to indicate cross-sections estimated from the cut factorization method and/or the weighting procedure.*

The contribution of the diboson events (WW/ZZ/WZ) has not been included in this study. From the results of ll and lh-channels studies [76] we assume it to be negligible.

In Figure 5.17 the final mass plot is shown for an integrated luminosity of 30 fb^{-1}. The background shape for $t\bar{t}$, $W + jets$ and QCD dijet has been obtained at the p_T^{total} cut level. Then it has been scaled to have the predicted number of events within the mass window.

For a conjectured Higgs boson with a given mass, m_H, the mass window has been defined as $m_H - 15$ GeV $\leq m_{\tau\tau} \leq m_H + 20$ GeV. A simple approach to estimate the expected significance of the signal is to count the events in this mass range and to compute $S/\sqrt{B}$, *i.e.* the probability of these events based on the background-only prediction. Table 5.12 shows the significance obtained with this procedure.

These results can be compared with those ones obtained for the ll and lh-channels. Table 5.13 shows the expected significance for different Higgs boson masses for the fully leptonic and semi-leptonic decay channels in 30 fb^{-1}. The significance has not been calculated counting the events, but fitting the $m_{\tau\tau}$ spectrum for signal and various backgrounds [76]. As for the hh-channel, the pileup is not included. Relying on a very high rejection against QCD jets given by the offline reconstruction algorithms for leptons, the background from QCD dijets has not been considered in that analysis. The final mass plots for these two decay channels are shown in Figure 5.18.

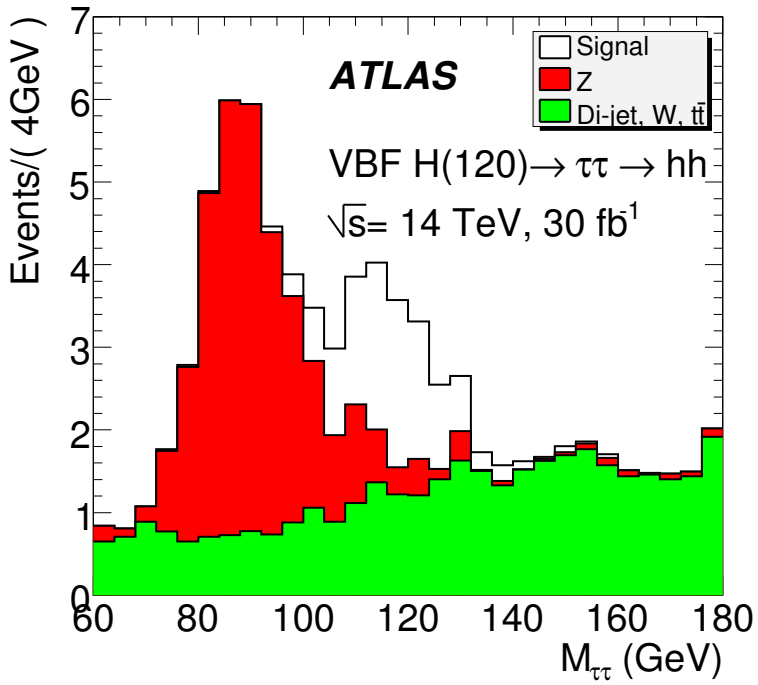

Figure 5.17: *Final mass plot for signal (m_H= 120 GeV) and background for an integrated luminosity of 30 fb^{-1}.*

m_H	Significance
105	1.6 ± 0.4
110	2.7 ± 0.8
115	2.6 ± 0.6
120	2.8 ± 0.6
125	2.4 ± 0.7
130	2.3 ± 0.6
135	1.6 ± 0.5

Table 5.12: *Expected signal significance for several masses based on number counting in a mass window with 30 fb^{-1} of data. Results are shown neglecting uncertainty in the background.*

m_H	ll-channel	lh-channel	combined
105	1.95	2.41	3.10
110	2.44	3.35	4.15
115	2.98	4.07	5.04
120	2.92	3.87	4.85
125	2.75	3.75	4.65
130	2.46	3.38	4.18
135	2.21	3.32	3.99
140	1.80	2.70	3.24

Table 5.13: *Expected signal significance for several masses based on fitting the $m_{\tau\tau}$ spectrum with 30 fb^{-1} of data.*

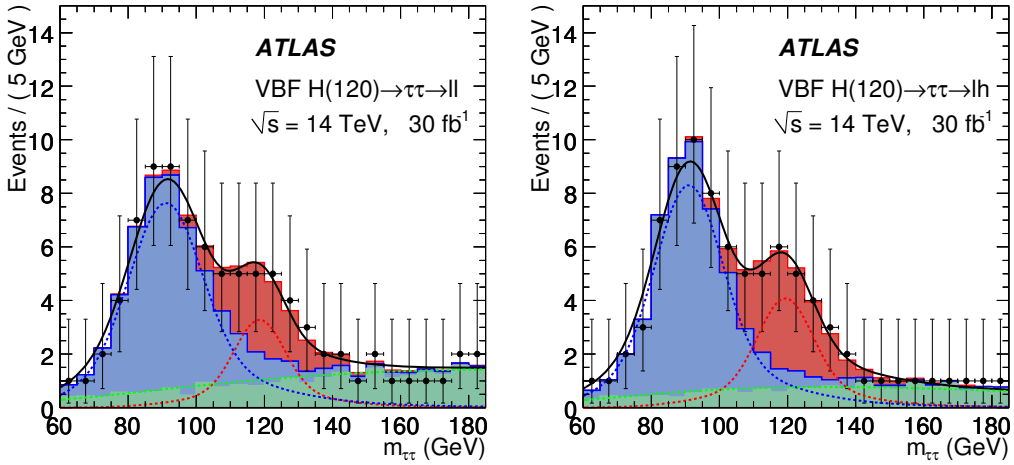

Figure 5.18: *Final mass plot for signal and background for ll-channel (left) and for lh-channel (right) at $m_H = 120$ GeV with 30 fb^{-1} of data. The fits are performed to the signal and background expectation (histograms), while the overlaid data with error bars are only indicative of a possible data set.*

5.4.5 Background Estimation from Data

The procedures that have been used to estimate the backgrounds have large uncertainties, especially the one from QCD multijets. For a pure QCD process to satisfy the event selection, there must be at least four high-p_T jets. Previous studies have shown that the parton shower underestimates the tagging jet requirement by a factor of 2$\sim$3 [27]. Thus a reliable estimation of QCD multijets must be made from data. A few handles exist for estimating the pure QCD background. First, one can utilize a sample of same-sign tau candidates to estimate the fake tau contribution since the sign of the tau candidate from QCD is approximately random. Potentially, one can utilize constraints from the track multiplicity distribution[76]. Furthermore, one can loosen the identification requirements on the tau candidates to obtain a sample dominated by QCD fakes, and then extrapolate this background into the signal region using knowledge of the fake tau's likelihood distribution obtained with data. The real effectiveness of these techniques needs also to be established with the data, but here we assume that the same-sign sample estimates the QCD background with uncertainties deriving by two factors. The first factor is the statistical error in the control sample, which scales like $1/\sqrt{N_{SS}}$, where N_{SS} is the size of same sign sample coming from QCD backgrounds. The second factor is the systematic uncertainty associated with the use of the same sign sample to estimate the sample with opposite signs. Experience from the Tevatron shows that the charge correlations can be of order 13% with an uncertainty of order 3% [85, 86]. Given that the final state requires two additional jets, which can alter the contribution of quark and gluon-initiated jets, we assume a 10% systematic error associated with the charge correlation. In addition, the same sign sample will also include a contribution from $W \to \tau\nu$+jets, where the Tevatron experiments observed a charge correlation that was much higher with a 40% uncertainty [85]. The experience from the Tevatron provides some insight for this approach, but the results are not directly relevant because the LHC is not a $p - \bar{p}$ machine.

Other data-driven background estimation procedures are under investigation to estimate the shape and the normalization of $Z + jets$, $W + jets$ and $t\bar{t}$ backgrounds [76]. The most serious danger of $Z + jets$ background comes from the high-side tail in the $m_{\tau\tau}$ distribution, where we would expect to see the signal. To model the $Z \to \tau\tau + jets$ background, $Z \to \mu\mu + jets$ events could be used in two different ways [87]. In the first way, after applying a minimal set of event selection criteria to the data, the measured momenta of the two lepton from the $Z \to \mu\mu$+jets events are fed into the TAUOLA decay package [69] and the tau decay products are passed to the ATLAS detector simulation and reconstruction software. In the second way a pure $Z \to \mu\mu$

sample with less than 1% contamination from the other processes is selected and then the muon momenta are rescaled to make them look like if coming from the process $Z \to \tau\tau \to \mu\mu + 4\nu$.

All the cuts are applied to these manipulated samples and the $m_{\tau\tau}$ invariant mass is reconstructed using the collinear approximation. The normalization of the $Z \to \tau\tau$ background does not require this emulation because it can be estimated directly from the height of the Z-peak in $m_{\tau\tau}$ spectrum obtained with the signal candidates. Figure 5.19 (left) shows the reconstructed visible mass for the true and emulated samples and (on the right) the bin-by-bin ratio of these distributions obtained applying the first procedure. Excellent agreement between the true and emulated distributions are observed in all the decay channels in the region of interest. The grey horizontal band in the figure represents $\pm 10\%$ around a ratio of 1, which is used to describe the uncertainty in the shape from the tau modeling and sensitivity to the analysis cuts. With this method we are able to accurately model the shape of the $Z \to \tau\tau$ backgrounds for all tau decays.

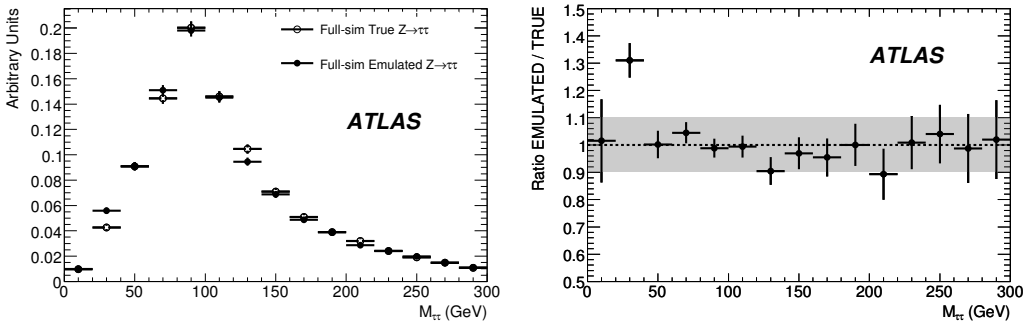

Figure 5.19: *Reconstructed invariant mass distribution (left) and its bin-by-bin ratio (right) generated from the true and emulated $Z \to \tau\tau$ events. The grey band represents $\pm 10\%$ around a ratio of 1.*

To estimate the contribution of fake τs from leptons and jets in $t\bar{t}$, $W + jets$, the track multiplicity around the τ candidate can be used. Real taus typically have one or three tracks, with some spread due to tracking efficiency or the presence of spurious tracks. Electrons have dominantly a single track, while jets have a broad distribution with a higher average multiplicity. Figure 5.20 shows the track multiplicity distribution for taus, electrons, and jets in a cone of radius 0.7 after removing outlying tracks [88]. It is clear that the track multiplicity can be used to constrain the relative abundance of the three components to the distribution. While Figure 5.20 was created from Monte Carlo, the electron and jet track multiplicity distributions can easily be obtained from data. Given a sample of tau candidates, the rela-

tive abundance of taus, electrons, and jets can be found by fitting the track multiplicity distribution.

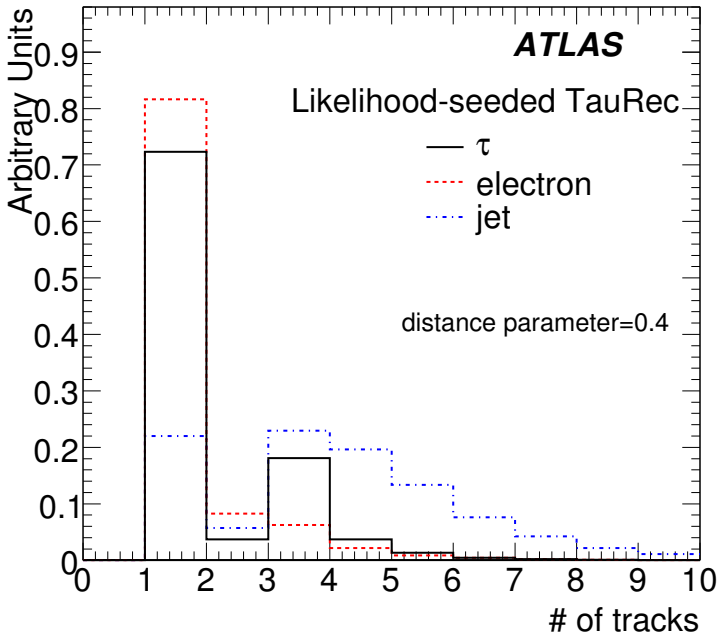

Figure 5.20: *Track multiplicity distribution for QCD fake events and electron-fake events as well as the τ signal.*

5.5 Systematics

The measurement of the Higgs boson mass is sensitive to the energy scale of electrons, muons, hadronic taus and E_T^{miss}. Furthermore, measurement and exclusion of $\sigma(pp \rightarrow qqH) \times BR(H \rightarrow \tau\tau)$ are sensitive to the uncertainty on the signal selection efficiency. Below we discuss the impact of these systematics on the analysis [76].

5.5.1 Systematic Mis-measurement of the Signal

First, we consider the purely experimental sources of systematics. The approach here is to assume that once we have data, the estimations of the energy scales, resolutions, and efficiencies might be systematically biased. Estimates of the systematic uncertainty from various sources are given in Table 5.14. There is not yet a dedicated study of the expected uncertainty

on these efficiencies from data, thus we assume uncertainty on the reconstruction efficiency to be half the tau identification efficiency and point out that most of the effects relevant for the jets have already been included in the jet energy scale uncertainty. We use the nominal detector performance as the central value and then manipulate the Monte Carlo signal to reflect these changes. For instance, in the case of the τ energy scale uncertainty, we coherently change all τs to have 5% higher $\mathrm{E_T}$, modify the $\mathrm{E_T^{miss}}$ vector accordingly, and recalculate the signal efficiency. This is done individually for each source of systematic and upward and downward fluctuations are treated separately. In the case of the jet energy scale, only some elements of the uncertainty are relevant for $\mathrm{E_T^{miss}}$. A study of $\mathrm{E_T^{miss}}$ projected onto the direction of the reconstructed Z in $Z \rightarrow ll$+jets with a subset of the analysis cuts indicated that the $\mathrm{E_T^{miss}}$ scale can be measured within 5%; thus we only manipulate the $\mathrm{E_T^{miss}}$ vector according to a 5% jet energy scale shift. In the case of systematic uncertainty on the resolution, we only considered a degradation in the resolution by the tabulated amount. Finally, for systematics on reconstruction and identification efficiency we assume a 1-to-1 transfer to the uncertainty on the signal efficiency and include a factor of two when the signal efficiency scales as the square (e.g. the tau efficiency in the hh-channel). Table 5.14 summarizes the effect of systematic mis-measurement on the signal efficiency. As we can see the jet energy scale is the dominant systematic effect.

Source	Relative uncertainty	Effect on signal efficiency
luminosity	$\pm 3\%$	$\pm$ 3%
tau energy scale	$\pm$ 5%	$\pm$ 4.9%
tau energy resolution	$\sigma(E) \oplus 0.45\sqrt{E}$	$\pm$ 1.5%
tau ID efficiency	$\pm$ 5%	$\pm$ 10%
jet energy scale†	$\pm$ 7% ($\lvert\eta\rvert \leq 3.2$) $\pm$ 15% ($\lvert\eta\rvert \geq 3.2$) $\pm$ 5% (on $\mathrm{E_T^{miss}}$)	$^{+16\%}/_{-20\%}$
jet energy resolution	$\sigma(E) \oplus 0.45\sqrt{E}$ ($\lvert\eta\rvert \leq 3.2$) $\sigma(E) \oplus 0.67\sqrt{E}$ ($\lvert\eta\rvert \geq 3.2$)	$\pm$ 1%
forward tagging efficiency	$\pm$ 2 %	$\pm$ 2%
central jet reconstruction efficiency	$\pm$ 2 %	$\pm$ 2%
total summed in quadrature		$\pm$20%

Table 5.14: *Estimated scale of systematic mis-measurements and their effect on the signal efficiency.* † *When varying the jet energy scale, only a 5% mis-measurement of the jet energy was used in manipulating the* $\mathrm{E_T^{miss}}$ *vector. See text for details.*

5.5.2 Theoretical Uncertainty

In addition to the effect of systematic mis-measurement on the signal efficiency, theoretical uncertainties also limit our ability to estimate the signal efficiency. Next-to-leading order QCD calculations are now available for the vector boson fusion process. A dedicated study [89] investigated the overall renormalization and factorization scale dependence (2%) as well as the parton distribution function (PDF) uncertainties (3.5%). Next-to-leading order electroweak corrections are also quite large for the vector boson fusion process, giving a 3% uncertainty for the full next-to-leading order calculation [90]. Recently, the dominant next-to-leading order QCD corrections to the Higgs boson plus three jets have been calculated for vector boson fusion, providing a scale uncertainty on the parton-level central jet veto survival probability of 1% [91].

While the parton-level theoretical uncertainties are under very good control and below the level of both the statistical error and measurement-related systematics, the same is not true for the theoretical uncertainty related to the parton-shower and underlying-event. We rely on Monte Carlo simulations that model the parton-shower, hadronization, and underlying event to simulate the detector response. The uncertainty in these calculations is not comparable to the accuracy of the parton-level predictions. The central jet veto efficiency was studied with the signal process generated with PYTHIA (with various tunings), HERWIG and SHERPA and the fast detector simulation. After the analysis cuts, the different generators differ by 41% [83]. Studies focusing specifically on the matrix element–parton shower matching indicate a substantially smaller uncertainty [92, 93]. We will measure the underlying event [94, 95] and tune the parton shower and hadronization with data, but it is likely that this contribution of the uncertainty will remain significant. Currently there is no estimate of the expected uncertainty related to the parton-shower, hadronization, and underlying event tuning. Clearly, this is an area that deserves attention as such a large uncertainty will hinder exclusions if a Higgs boson does not exist in this mass range and cross-section and coupling measurements if it does. After discussions with the authors of PYTHIA, HERWIG, and SHERPA we feel that the residual uncertainty in the parton shower after tuning to the data will be less than the 18% uncertainty quoted for the jet energy scale. Thus, the uncertainty in the signal efficiency will be dominated by the jet energy / E_T^{miss} scale uncertainty, and the precise uncertainty in the parton shower is less relevant. Table 5.15 summarizes the theoretical uncertainties for the signal production.

Source	Relative uncertainty	Effect on signal efficiency
PDF uncertainties	$\pm$3.5%	$\pm$3.5%
scale dependence on cross-section	$\pm$3%	$\pm$ 3%
scale dependence CJV efficiency	$\pm$ 1%	$\pm$ 1%
parton-shower and underlying event	$\pm \leq$10%	$\pm <$10%
total summed in quadrature		$\pm <$ 10%

Table 5.15: *Theoretical uncertainties which affect the estimation of the signal efficiency.*

5.6 Multijet Background Estimation with the First Data

The first measurements that will be done at LHC will be related to jet physics. For this analysis it will be fundamental to understand QCD events with a very peculiar kinematic configuration. These events must have at least four reconstructed jets, two of them in the central region with p_T >35–30 GeV, the other two in opposite η sides of the detector, with a large η separation, high p_T and a large invariant mass. Moreover events will be required to have a E_T^{miss} greater than 40 GeV. To estimate the amount of the events that will be recorded by ATLAS in the first inverse picobarns of data, we have used three ATLFAST samples of QCD multijets (3, 4 and 5jets) calculated by ALPGEN [30] and processed trough HERWIG [64] for fragmentation, hadronization and parton shower. These events were required to have:

- $E_T^{miss} > 40$ GeV;
- at least four jets;
- two jets within $|\eta| < 2.5$ with $p_T > 35$ and 30 GeV;
- two jets with $p_T > 40$ and 20 GeV, $\eta_{j_1} \times \eta_{j_2} < 0$, $\Delta\eta > 4$, $m_{j_1 j_2} > 700$ GeV.

Table 5.16 shows the cross section for each sample, the integrated luminosity corresponding to the available statistics of the sample itself, and the result of the cut flow.

Considering an integrated luminosity of 10 pb^{-1}, it will be possible to collect at least one thousand events with the required configuration, if they are triggered. ATLAS trigger foresees single jet triggers, multijet triggers and jet trigger in association with E_T^{miss}. Two of those ones that are supposed to run at the luminosity of $10^{31} cm^{-2} s^{-1}$ are `2j42_xe30` and `j70_xe30`. They require a $E_T^{miss} > 30$ GeV and two jets with $p_T > 42$ GeV or one jet with

Sample	QCD 3jets	QCD 4jets	QCD 5jets
Cross section (pb)	17×10^6	2.63×10^6	5.21×10^5
Integrated luminosity(pb^{-1})	0.074	0.87	3.84
Initial events	1255000	2295000	2000000
$E_T^{miss} > 40$ GeV	820	2626	5053
4 jets	385	2145	4999
2 central jets	131	1275	4081
2 forward	3	23	164
`2j42_xe30` OR `j70_xe30`	3	20	157

Table 5.16: *Top: cross section of three different multijet samples, with the integrated luminosity corresponding to the sample statistics. Middle: Results of the cut flow. Bottom: number of events having the offline objects requested by the triggers* `2j42_xe30` *or* `j70_xe30`*.*

$p_T > 70$ GeV respectively. Since in ATLFAST no information about the trigger decision is available, we verified that the surviving events had the reconstructed objects requested by the trigger. The ATLAS jet trigger cover the eta region -3.2 $< \eta <$ 3.2, thus the reconstructed jets have to lie in this region. As it is shown in the last line of Table 5.16, almost all the events have the reconstructed configuration that should be accepted by the two chosen triggers.

Thus it seems possible to have already a few thousand events in a few tens of pb^{-1}, that would allow to study the shape and the normalization of the multijet background for our Higgs channel.

Conclusions

The discovery potential for Standard Model Higgs boson in the mass range from 105-135 GeV has been studied in the vector boson fusion production mode with subsequent $H \rightarrow \tau^+\tau^-$ decay at the ATLAS detector. Fully hadronic decay of a τ-pair has been investigated for the first time, using the most recent Monte Carlo generators, a full GEANT-based simulation of the ATLAS detector and the current trigger and reconstruction algorithms. Being a fully hadronic final state, its feasibility in the environment of the hadronic machine of the LHC is very challenging.

It requires a dedicated tau trigger with very good performance both in term of efficiency and rate. It can be used in association with E_T^{miss} trigger or with another tau trigger. The detector characteristics are also fundamental to obtain a very high jet rejection at offline level together with a precise measurement of the $\mathrm{E_T^{miss}}$. The algorithms used by the ATLAS detector both for the trigger and offline reconstruction have been discussed in this work.

The peculiar signature of the vector boson fusion production mechanism can be exploited to further reduce the background due to QCD multijets, $Z(\rightarrow \tau\tau) + jets$, $W(\rightarrow \tau\nu) + jets$, $t\bar{t}$.
Because of the limited size of the Monte Carlo background samples, factorization and weighting procedures have been applied to estimate the final sensitivity. ATLFAST samples have been used to give an estimation of the multijet QCD background, thus a brief discussion on the fast simulation principles and on its limits has been included. We have verified that the uncertainties related to these procedures will not disturb the analysis once we have data. The results of the fully hadronic tau final state are very similar to those ones of the ll- and lh-channels for the signal and non-QCD background. The big uncertainty on the QCD multijet background will be removed once we have data. In fact we estimated that in few tens of inverse picobarns of recorded data at L= $10^{31}\mathrm{cm}^{-2}\mathrm{s}^{-1}$, we can collect some thousands of multijet events with a adequate kinematic configuration to allow the study of the shape and the normalization of this important background.

Methods to determine the background shape from data have been addressed. An evaluation of the major theoretical and experimental systematics has been done.

We estimate that signal significance of 2-3 can be obtained in the mass range from 105-135 GeV, without including systematics and not considering the pileup.

In spite of the large uncertainties in the background estimation on the search of the VBF Higgs boson decaying in fully hadronic τ-pairs, due to the importance of the assessment of Higgs existence in this difficult mass range, we believe that it is crucial to continue this analysis which here we have shown can realistically be performed.

The results coming from the fully hadronic τ-pair decay will be eventually combined with the purely leptonic and semileptonic decays, improving the discovery potential of the Higgs boson which is the major aim of the LHC.

Acknowledgments

People I wish to thank for having contributed to this work, directly and not, are really many.

First of all Prof. Vincenzo Cavasinni, for his frequent and useful suggestions and for the time he spent to discuss with me and with the people of the ATLAS Higgs Working Group. Inside this group I want to thank in particular Markus Schumacher, Louis Fayard and Aleandro Nisati for their support, Donatella Cavalli for the useful discussions and the precious hints, Rachid Mazini and Stefania Xella, with whom it has been a big pleasure to work.

Many thanks to the referees of this thesis, Peter Loch, Mauro Dell'Orso and Giovanni Carboni, who were extremely kind and ready to help. I would like to thank Peter Loch for his careful help for the final version of the thesis, Mauro dell'Orso for the precious discussions I have had with him, Giovanni Carboni for his comments and for the attention he has given to this work.

I want to thank all the people of the Pisa ATLAS group, in particular Prof. Tarcisio Del Prete and Chiara Roda, always ready to help me during these years, and Iacopo Vivarelli and Andrea Dotti for all the pleasant time spent together working and not working.

Many thanks to all the people of the TileCal collaboration, in particular Giulio, Belen, Esteban, Gerolf, Cristobal, David and Giorgi who have made pleasant the work during the testbeam and the commissioning periods.

I also wish to thank the friends of INFN with whom I shared courses, lunches, dinners and a lot of fun, and the friends at CERN, always ready to give me a house, food, and their very good company. Therefore, many thanks to Ersilio, Ans, Robbo, Maiko, Jose, Beppe, Alex, and to all the other friends I cannot mention here because they are too many.

A special thanks is to Claudio, for having encouraged me during these last years and for his patience in reading this thesis.

Bibliography

[1] S. L. Glashow, Nucl. Phys. **20**, 579 (1961).

[2] S. Weinberg, Phys. Rev. Lett. **19**, 1264 (1967).

[3] A. Salam, *Elementary Particle Theory*, N. Svartholm, Stockholm, 1968.

[4] P. W. Higgs, Phys. Rev. Lett. **12**, 132 (1964).

[5] P. W. Higgs, Phys. Rev. Lett. **13**, 508 (1964).

[6] P. W. Higgs, Phys. Rev. Lett. **145**, 1556 (1966).

[7] The LEP Electroweak Working Group , CERN-PH-EP/2004-069 (2004).

[8] E. Costa *et al.*, Nucl. Phys. **B297**, 244 (1988).

[9] S. Fanchiotti *et al.*, Phys. Rev. , 307 (1993).

[10] W. J. Marciano *et al.*, Phys. Rev. Lett. **65**, 2963 (1990).

[11] S. Eidelman *et al.*, Phys. Lett. **B592**, 1 (2004).

[12] B.W.Lee *et al.*, Phys. Rev. Lett. **38**, 883 (1977).

[13] M.Spira and P.M.Zerwas, hep-ex/9803257 (1997).

[14] J. Ellis *et al.*, Nucl. Phys. **B106**, 292 (1976).

[15] The LEP Working Group for Higgs Boson Searches, Phys. Lett. **B565**, 61 (2005).

[16] The LEP Electroweak Working Group , CERN-PH-EP/2007-039 (2007).

[17] http://lepewwg.web.cern.ch/LEPEWWG/

[18] The CDF II Collaboration, *The CDF II Technical Design Report*, FERMILAB-Pub-96/390-E, 1996.

[19] T. LeCompte *et al.*, Ann. Rev. Nucl. Part. Sci. **50**, 71 (2000).

[20] The TEVNPH Working Group, FERMILAB-PUB-08-069-E (2008)

[21] R.M. Roser, *The Tevatron and the CDF Experiment: A Year in Review*, Colloquium presentation given at Fermilab on January 16, 2008.

[22] http://www-cdfonline.fnal.gov/ops/opshelp/stores/kumac/store_tot.gif

[23] G. Bernardi, *Searches and Prospects for Standard Model Higgs boson at the Tevatron*, Talk given at the Hadron Collider Physics Symposium (HCP2008)

[24] M. Herndon, *Searches for the Higgs Boson*, talk given at the 34^{th} International Conference on High Energy Physics, ICHEP 08.

[25] ATLAS Collaboration, *Detector and Physics Performance Technical Design Report*, CERN/LHCC/99-15, 1999.

[26] D. Rainwater and D.Zeppenfeld, Phys. Rev. **D60**, 113004 (1999).

[27] S. Asai *et al.*, Eur. Phys. J. **C32**, 19 (2003).

[28] V. Cavasinni, F. Sarri, I. Vivarelli, R. Mazini, *A search for VBF Higgs production with double hadronic tau decays in ATLAS*, ATL-COM-PHYS-2008-076.

[29] T. Sjöstrand, hep-ph/0108264 (2002).

[30] M.L. Mangano, *et al.*, J. High Energy Physics **07** (2003) 001.

[31] E. Boos *et al.*, hep-ph/9908288 (1998).

[32] L.Evans and P.Bryant, *LHC Machine*, 2008 JINST **3** S08001 (2008).

[33] CMS Collaboration, *CMS Technical Proposal*, CERN/LHCC/94-38, 1994.

[34] ALICE Collaboration, *ALICE Technical Design Report*, CERN/LHCC/01-021, 2001.

[35] LHCb Collaboration, *LHCb Technical Proposal*, CERN/LHCC/98-004, 1998.

[36] ATLAS Collaboration, *ATLAS Technical Proposal*, CERN/LHCC/94-43, 1994.

[37] ATLAS Collaboration, *ATLAS Magnet System Technical Design Report*, CERN/LHCC/97-18, 1997.

[38] ATLAS Collaboration, *ATLAS Inner Detector Technical Design Report*, CERN/LHCC/97-16, 1997.

[39] ATLAS Collaboration, *ATLAS Tile Calorimeter Technical Design Report*, CERN/LHCC/96-42, 1996.

[40] ATLAS Collaboration, *Detector and Physics Performance Technical Design Report*, CERN/LHCC/99-14, 1999.

[41] ATLAS Collaboration, *ATLAS Liquid Argon Calorimeter Technical Design Report*, CERN/LHCC/96-41, 1996.

[42] W. H. Press *et al.*, *Numerical Recipes in C*, chapter 13, Cambridge University Press, 2002.

[43] B. Dowler *et al.*, Nucl. Inst. and Meth. **A482**, 94 (2002).

[44] http://hep.uchicago.edu/atlas/electr/writeups/rome_paper.pdf.

[45] ATLAS Collaboration, *ATLAS Muon Spectrometer Technical Design Report*, CERN/LHCC/97-22, 1997.

[46] ATLAS Collaboration, *ATLAS High-Level Trigger, Data Acquisition and Controls Technical Design Report.* CERN/LHCC/03-22, CERN, 2003.

[47] ATLAS Collaboration, *ATLAS Computing Technical Design Report.* CERN/LHCC/05-22, CERN, 2005.

[48] M. Aderholz *et al.*, *Models of Networked Analysis at Regional Centres for LHC Experiments (MONARC), Phase 2 Report*. CERN/LCB/00-01, CERN, 2000.

[49] B.Caron, R.Soluk, J.Pinfold, *A Study of the Tau/Hadron Trigger*, ATL-COM-DAQ-2003-030.

[50] J.Garvey *et al.*, *Use of a FPGA to identify electromagnetic clusters and isolated hadrons in the ATLAS Level-1 Calorimeter trigger*, Nuclear Instruments and Methods A,vol. 512 no. 3, 2003, pp.506-516.

[51] ATLAS Collaboration, *ATLAS First Level Trigger Technichal Design Report*, CERN/LHCC/98-14, 1998.

[52] ATLAS Collaboration, *ATLAS Tau Trigger* in *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*,
CERN-OPEN-2008-020, Geneva, 2008, and reference therein, arXiv:0901.0512v1 [hep-ex].

[53] ATLAS Collaboration, *Reconstruction and Identification of Hadronic τ Decay with ATLAS* in *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*,
CERN-OPEN-2008-020, Geneva, 2008, and reference therein, arXiv:0901.0512v1 [hep-ex].

[54] The ATLAS Collaboration, *Detector Level Jet Corrections* in *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*,
CERN-OPEN-2008-020, Geneva, 2008, and reference therein, arXiv:0901.0512v1 [hep-ex].

[55] ATLAS L1Calo Group, *ATLAS Level-1 Calorimeter Trigger Algorithms*, ATL-DAQ-2004-011.

[56] The ATLAS Collaboration, *Triggering top quark events in ATLAS* in *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*,
CERN-OPEN-2008-020, Geneva, 2008, and reference therein, arXiv:0901.0512v1 [hep-ex].

[57] The ATLAS Collaboration, *Measurement of Missing Transverse Energy in ATLAS* in *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*,
CERN-OPEN-2008-020, Geneva, 2008, and reference therein, arXiv:0901.0512v1 [hep-ex].

[58] C. Amsler *et al.*,Physics Letters **B667**,1 (2008).

[59] M. Heldmann and D. Cavalli, *An improved tau-Identification for the ATLAS experiment*, ATLAS Note, ATL-PHYS-PUB-2006-008

[60] C. Cojocaru et al.,*Hadronic calibration of the ATLAS liquid argon endcap calorimeter in the pseudorapidity region 1.6-1.8 in beam tests*, Nucl. Inst. and Meth.**A531**, 481-514 (2004).

[61] ATLAS Collaboration, *Performance of Calorimeter Clustering Algorithms* in *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*,
CERN-OPEN-2008-020, Geneva, 2008, arXiv:0901.0512v1 [hep-ex].

[62] R. K. Ellis *et al.*, Higgs decay to $\tau^+\tau^-$: A possible signature of Intermediate Higgs Bosons at the SSC, Nucl. Phys. B297 (1988) 221.

[63] L. DiLella, Proceedings of the Large Collider Workshop, edited by G. Jarlskog and D. Rein (Aachen, 4-9 October 1990) CERN 90-10/ECFA 90-133, Vol. II, p. 530.

[64] G. Marchesini and B. R. Webber, Nucl. Phys. **B310** (1988) 461;
G. Marchesini et al., Comput. Phys. Commun. **67** (1992) 465.

[65] M.L. Mangano, *Exploring theoretical systematics in the ME-to-shower MC merging for multijet process* in Proceedings of Matrix Element/Monte Carlo Tuning Workshop, Fermilab, Nov. 15, 2002.

[66] T. Gleisberg, S. Hoche, F. Krauss, A. Schalicke, S. Schumann and J. C. Winter, JHEP **0402** (2004) 056 [arXiv:hep-ph/0311263].

[67] S. Frixione and B.R. Webber, J.High Energy Phys. **06** (2002) 029.

[68] M.A. Dobbs *at al.*, hep-ph/0403045.

[69] S. Jadach et al., Comput. Phys. Commun. **76** (1993) 361.

[70] E. Barberio, Z. Was, *PHOTOS: A universal Monte Carlo for QED radiative corrections: version 2.0* , Comput. Phys. Commun. **79** (1994) 291.

[71] B.Mallado et al., *Higgs Production Cross-Sections and Branching Ratios for the ATLAS Higgs Working Group*,
ATL-COM-PHYS-2007-024 (2007).

[72] E. Richter-Was *et al.*, *ATLFAST 2.0 a fast simulation package for ATLAS*
ATL-PHYS-98-131 (1998).

[73] ATLAS Collaboration, *Monte Carlo Generators for the ATLAS Computing System Commissioning*, in *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*,
CERN-OPEN-2008-020, Geneva, 2008, arXiv:0901.0512v1 [hep-ex].

[74] `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/ZplusJetsAlpgen`

[75] `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/WplusJetsAlpgen`

[76] A. Ahmad *et al.*, *Search for the Standard Model Higgs boson via Vector Boson Fusion production process in the di-tau channels with ATLAS*, ATL-PHYS-INT-2008-0XX, arXiv:0901.0512v1 [hep-ex].

[77] D. L. Rainwater, R. Szalapski and D. Zeppenfeld, *Probing color-singlet exchange in Z + 2-jet events at the LHC*, Phys.Rev.**D54**, 6680 (1996) [arXiv:hep-ph/9605444].

[78] ATLAS Collaboration, *Jet Reconstruction Performance*, in *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*, CERN-OPEN-2008-020, Geneva, 2008, arXiv:0901.0512v1 [hep-ex].

[79] M. Campanelli *et al.*, *A study of jet algorithms using the SpartyJet tool*, ATLAS-COM-PHYS-2008-021.

[80] I. Tsukerman, *Forward jet tagging and central jet veto in VBF Higgs CSC MC samples*, ATL-COM-PHYS-2007-083.

[81] I. Tsukerman, *Characteristics of jets and E_T^{miss} in VBF Higgs and minimum-bias CSC MC samples*, ATL-PHYS-INT-2008-009.

[82] C. Ruwiedel, M. Schumacher, and N. Wermes, *Prospects for the Measurement of the Structure of the Coupling of a Higgs Boson to Weak Gauge Bosons in Weak Boson Fusion with the ATLAS Detector*, Eur. Phys. J. **C51**:385–414, (2007). [SN-ATLAS-2007-060]

[83] M. Röder *et al.*, *Different Monte Carlo Generators in Higgs VBF Analyses and CJV optimization*, ATL-COM-PHYS-2008-025

[84] D. Cavalli *et al.*, *Performance of the ATLAS fast simulation ATLFAST*, ATL-PHYS-INT-2007-005 (2007).

[85] V. M. Abazov *et al.* [D0 Collaboration], *First measurement of $\sigma(\bar{p}p \to Z)Br(Z \to \tau\tau)$ at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. D **71** (2005) 072004. Further discussion in preliminary draft of *Measurement of $\sigma(\bar{p}p \to Z)Br(Z \to \tau\tau)$ with 1 fb^{-1} at $\sqrt{s} = 1.96$ TeV* found here: http://www-d0.fnal.gov/Run2Physics/WWW/results/prelim/EW/E21/E21.pdf

[86] S. Duensing, *Measurement of $\sigma(\bar{p}p \to Z)Br(Z \to \tau\tau)$ at $\sqrt{s} = 1.96$-TeV using the D0 detector at the Tevatron*, FERMILAB-THESIS-2004-23

[87] X. Chen *et al.*, *A Data-driven Controll Sample for $Z \to \tau^+\tau^-$+jets Backgrounds to Higgs Searches*, ATL-COM-PHYS-2008-144.

[88] S. Tsuno and J. Tanaka, *$Z \to \tau\tau \to lh$ measurement using early data in ATLAS*, ATL-COM-PHYS-2007-050.

[89] T. Figy, C. Oleari and D. Zeppenfeld, *Next-to-leading order jet distributions for Higgs boson production via weak-boson fusion*, Phys. Rev. D **68** (2003) 073005 [arXiv:hep-ph/0306109].

[90] M. Ciccolini, A. Denner and S. Dittmaier, *Strong and electroweak corrections to the production of Higgs+2jets via weak interactions at the LHC*, Phys. Rev. Lett. **99** (2007) 161803 [arXiv:0707.0381 [hep-ph]].

[91] T. Figy, V. Hankele and D. Zeppenfeld, *Next-to-leading order QCD corrections to Higgs plus three jet production in vector-boson fusion*, arXiv:0710.5621 [hep-ph].

[92] V. Del Duca *et al.*, *Monte Carlo studies of the jet activity in Higgs + 2jet events*, JHEP **0610** (2006) 016 [arXiv:hep-ph/0608158].

[93] J. Alwall *et al.*, *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. C **53** (2008) 473 [arXiv:0706.2569 [hep-ph]].

[94] ATLAS Collaboration, *A study of minimum bias at sqrt(s)=14 TeV*, in *Expected Performance of the ATLAS Experiment, Detector, Trigger and Physics*,
CERN-OPEN-2008-020, Geneva, 2008, and reference therein, arXiv:0901.0512v1 [hep-ex].

[95] D. Acosta, et al., The CDF Collaboration, *Phys. Rev.***D70**2004072002.